PROTEIN ENGINEERING STUDIES ON BACILLUS THERMCATENULATUS LIPASE

by

Emel Durmaz

Submitted to the Graduate School of Engineering and Natural Sciences
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Sabancı University

October 2012

APPROVED BY:

Assoc. Prof. Dr. Uğur Sezerman.....
(Thesis Supervisor)

Prof. Dr. Serdar Kuyucak.....

Prof. Dr. Selim Çetiner.....

Assoc. Prof. Dr. Levent Öztürk.....

Asst. Prof. Dr. Alpay Taralp.....

DATE OF APPROVAL: 15.10.2012

Protein Engineering Studies On *Bacillus thermocatenulatus* Lipase

Emel Durmaz
BIO, PhD Thesis, 2012
Thesis Supervisor: Uğur O. Sezerman

Keywords: Protein engineering, Lipase, Selectivity, Stability

Abstract

Bacillus thermocatenulatus lipase (BTL2) is a thermostable enzyme with distinct tributyrin selectivity and currently two different resolved structures are publicly available representing the; a) active-monomer form (PDB ID: 2w22) and b) inactive-dimer form (PDB ID: 1KU0). Using the information gained from these two structures, here we report the results of two studies related to the protein engineering of BTL2 aiming to improve its chain-length selectivity and to assess the impact of the lid tryptophans. For both of these studies, the experimental approach was to generate specific BTL2 mutants for which site-directed mutagenesis, heterologous protein expression and purification methodologies were used.In the first study, a rational design of BTL2 was carried out to lower the activation barrier of hydrolysis of short chain substrates. The results indicate that the mutant L360F optimizes the physical and chemical interactions in BTL2 to lower the activation barrier for C4 and elevate it for C8 hydrolysis. For the second case, a semi-rational design was implemented to investigate the impacts of the two tryptophans (W212, W235) in the lid of BTL2. W235A affected tributyrin selectivity while W212A mutation is resulted in loss of the thermostability and aggregation tendency but also in the improvement of zinc tolerance in BTL2. For this, it is envisioned that the aggregation of BTL2 is mediated through a zinc domain and W212 is a secondary but vital residue for this action.

Bacillus thermocatenulatus Lipazı Üzerine Protein Mühendisliği Çalışmaları

Emel Durmaz
BIO, Doktora Tezi, 2012
Tez Danışmanı: Uğur O. Sezerman

Anahtar Kelimeler: Protein Mühendisliği, Lipaz, Seçicilik, Stabilite

Özet

Bacillus thermocatenulatus lipazı (BTL2) yüksek tribütirin seçiciliği olan thermokararlı bir lipazdır ve yapısı aktif- monomer (PDB ID: 2w22) ve inaktif-dimer (PDB ID: 1KU0) olarak atomistik detayda çözülmüştür. Burada, BTL2'nin protein mühendisliği ile ilgili iki vaka çalışması rapor edilmiştir. Protein mühendisliğinin rutin metotlarından olan yönlendirilmiş mutagenez, heterolog protein ekpressiyonu ve purifikasyon, bu tezdeki BTL2 ve mutantları için yapılmıştır. İlk olarak, kısa zincir boyundaki sübstratlara karşı olan aktivasyon bariyerini düşürmek için BTL2'nin akılcı dizaynı gerçekleştirilmiştir. So- nuçlar L360F mutantının BTL2'deki aktivasyonu bariyerini C4 için düşüren ve C8 için yükselten bir mutasyon olduğunu doğrulamıştır. İkinci olarak, yarı-akılcı dizayn kullanarak BTL2'nin kapak kısmında yer alan iki triptofan'nın (W212, W235) etkilerini incelemek için uygunlanmıştır. Özellikle W212A, thermokararlılık ve agregasyon açısından ciddi bir düşüş ve çinko toleransı da bir artış göstermiştir. Bu açıdan, BTL2'nin agregasyonun çinko bağlanma bölgesi üzerinden gerçekleştiğini ve W212 rezidüsünün bu mekanizma için sekonder ama elzem bir rolu ol- duğu öngörülmektedir.

Acknowledgements

It is a pleasure to express my humble gratitude to several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this dissertation.

I gratefully thank to my thesis advisor Assoc. Dr. Uğur Sezerman, for his guidance support throughout thesis. I also wish to convey my gratitude to thesis jury members; Prof. Dr. Serdar Kuyucak for sharing his exceptional scientific background; Asst. Dr. Alpay Taralp for his comments on enzyme kinetics; Prof. Dr. Selim Çetiner and Assoc. Dr. Levent Öztürk for their constructive comments on this dissertation. I am grateful that in the midst of all their activity, they accepted to participate.

I also wish to acknowledge the financial support of TÜBİTAK-BİDEB.

I convey my sincere thanks to my friends: Ali Eken, Cem Meydan, Günseli Akçapınar, Tuğsan Tezil, Aydın Albayrak, Özgür Gül, Cağrı Bodur, Batuhan Yenilmez for their advice and their willingness to share their bright thoughts with me on every kind of subject; Hazal Yılmaz, Serkan Sırlı, Duygu Soysal for their assistance in the molecular biology experiments. Including the ones I forgot to mention, I thank to all friends and fellows for providing the necessary motivation to take the load off my shoulder.

Last but not least; I would like to thank Can and my parents for their support and being there when I needed them to be and to Seher for being a *deadringer* to me.

TABLE OF CONTENTS

List of Tables

List of Figures

List of Symbols and Abbreviations

ΔG_S	Substrate binding energy
$\Delta G^{\ddagger}$	Activation energy
$\Delta G_T^{\ddagger}$	Activation barrier
DLS	Dynamic Light Scattering
ES	Enzyme-subtrate Complex
$ES^{\ddagger}$	Transtion state complex
4MU-C4	4-Methylumbelliferyl butyrate
4MU-C8	4-Methylumbelliferyl caprylate
ANS	1-anilinonaphthalene-8-sulfonate
BTL2	*Bacillus thermocatenulatus* lipase 2
Cα	Carbon alpha atom
CD	Circular dichroism
C_P	Standard Deviation of Hydrodynamic Radius
HBI	Hydrogen bond I
HBII	Hydrogen bond II
k_{cat}	Enzyme turnover
K_M	Michealis-Menten Constant
K_S	Binding constat for enzyme substrate complex
MD	Molecular Dynamics
PDI	Polydispersity Index
R_H	Hydrodynamic Radius
RMSD	Root-mean-square distance
RPM	Revolutions per minute
SDS-PAGE	Sodium doedecyl sulphate-polyacrylamide gel electrophoresis
SMD	Steered Molecular Dynamics
trp	Tryptophan
V_0	Initial velocity of the enzymatic reaction
V_{max}	Maximum velocity of the enzymatic reaction

1 LIPASES

1.1 Background

Lipases are metabolic enzymes found in every domain of life. They are produced by plants, animals and microorganisms. The first discovery of lipases by Eijkmann dates back to 1900s. He observed that several bacteria can produce and secrete lipases to the extracellular environments to degrade lipids. Considering the large variety of microbes known to human today, microbial lipases are one of the most populated enzyme class and they have been extensively studied with respect to enzyme promiscuity [1, 2]. Lipases (EC 3.1.1.3, triacylglycerol lipase) belong to hydrolases and they have become the objects of many successful protein engineering approaches [3–10]. Additionally, a database has been established to source the information on lipases engineering (http://www.led.uni-stuttgart.de/) [11]. As lipases have been extensively studied, their mechanism of action, selectivity, structure and kinetics are already determined and such information could be very useful in the context of lipase engineering. Here, this chapter gives an overview of lipases.

1.2 Reactions

Lipases catalyze the breakdown of triacylglycerols (TAG) to free fatty acids (FA), diacylglycerols (DAG), monoacylglycerols (MAG) and glycerol. This hydrolysis is an equilibrium reaction thus it can be perturbed by alteration of the concentration of substrates or products. Water is one of the reactants of hydrolysis reaction and thus changing the hydrolytic conditions of the reaction shifts the equilibrium (e.g., low hydrolytic conditions favor ester synthesis) [12].

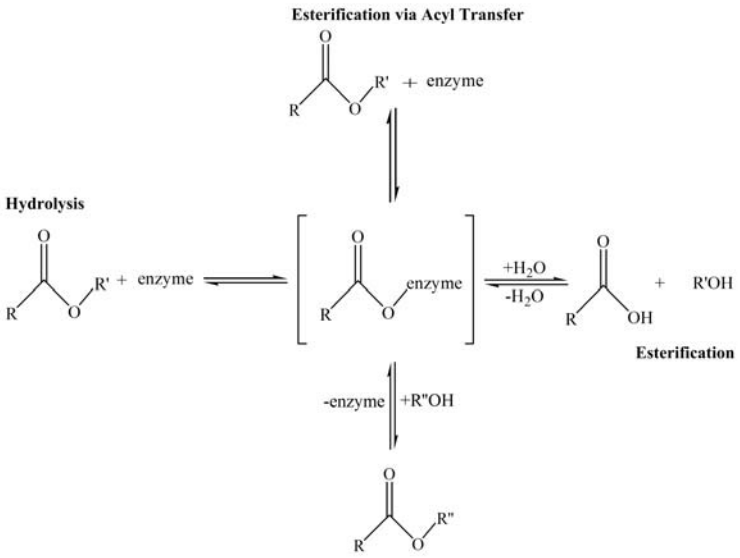

Figure 1.1:
Lipase-catalyzed reactions. Three main lipase reactions are shown and for all, first an acyl-enzyme intermediate is formed, then this unstable intermediate collapses to free enzyme and an acid in hydrolysis or an ester in esterification and transesterification. In aqueous media, the equilibrium is towards ester hydrolysis and in dry media it is towards ester synthesis. In organic solvents, lipase can catalyze acyl transfer reactions to synthesize new esters.

Lipases act on carboxylic ester bonds, either by breaking as in hydrolysis or by forming as in esterification and they also catalyze acyl transfer reactions through esterification. Depending on the chemical nature of the reactants and water availability in the media, different outcomes are possible yet, triacylglycerols (TAGs) are their natural substrates [13]. A carboxyl/thiolester or amide can also be formed under low water conditions. An ester group as an acyl donor can form the acyl-enzyme intermediate by either releasing an acid -acyl acceptor is water- or forming a new ester -acyl acceptor is alcohol/thiol/amine-. A summary of lipase-catalyzed reactions are shown in Figure 1.1. Acidolysis, alcoholysis, interesterification and aminolysis are the examples to transesterification which is shown as ”esterification via acyl transfer” in Figure 1.1.

1.3 Mechanism

Despite the versatility of the reactions catalyzed by lipases, their reaction mechanism is unique. The catalytic machinery is conserved in all lipases and is composed of three residues: serine, histidine and aspartate/glutamate -identical to serine-proteases- [14]. Figure 1.2 illustrates the reaction mechanism of lipases in triacylglycerol hydrolysis. The mechanism involves alignment of two residues (histidine and aspartate/glutamate) to lower the pKa of the serine hydroxyl which enables serine to carry out a nucleophilic attack on the ester bond.

According to Figure 1.2, the substrate which is an acyl donor contacts with the active site of the lipase to form the enzyme-substrate (ES) complex, or the Michaelis complex in honor of its founder. The histidine in the catalytic triad plays the role for a general base and activates the hydroxyl group of serine. Serine is ready to carry out a nucleophilic attack on the carbonyl carbon of the substrate to form the first tetrahedral intermediate. The negative charge on the oxyanion is stabilized in the hole generated by the main-chain amide groups of two residues and the positive charge on the histidine is stabilized by the aspartate or glutamate of the catalytic triad. The tetrahedral intermediate decomposes into first leaving group of the substrate, an alcohol, and the acyl enzyme intermediate.

After the formation of the acyl enzyme intermediate, a second tetrahedral intermediate is formed which corresponds to the highest energy barrier in the reaction, similar to the first intermediate (Figure 1.3). This intermediate is also collapsed to yield the deacylated-free form of the enzyme which is ready for the next cycle and the hydrolysis of the second substrate, an acid. During this deacylation step, a proton is transferred from the substrate to histidine via serine oxygen.

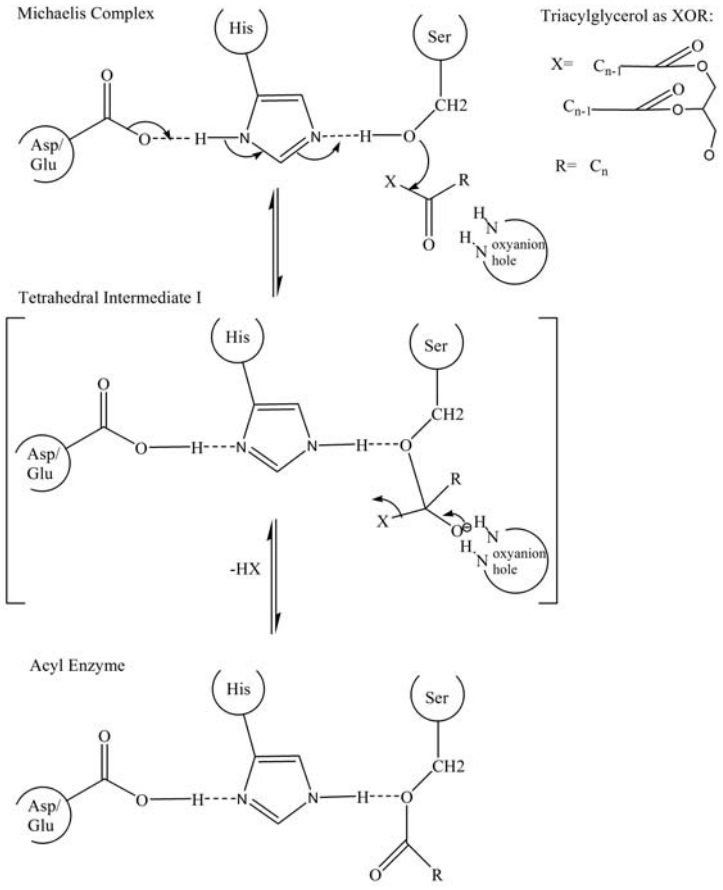

Figure 1.2: Mechanism of Acylation.

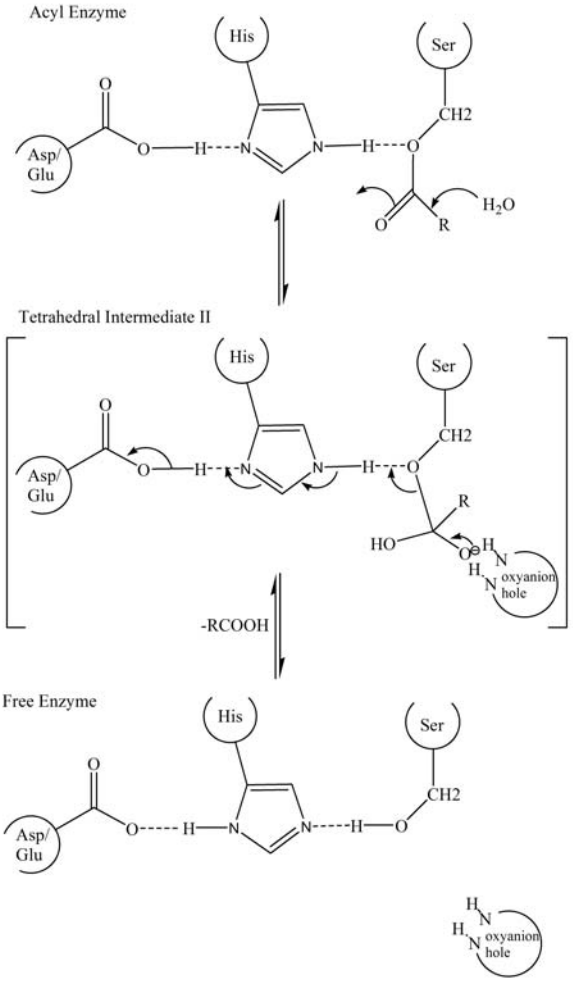

Figure 1.3: Mechanism of Hydrolysis and Deacylation.

1.4 Selectivity

Lipase selectivity towards TAGs is usually distinguished as regio (related to the position of the ester bond in TAG), stereo-selectivity (related to the chiral center) and substrate selectivity (related to the type and chain-length) [15, 16].

1.4.1 Regio-selectivity

Also referred as positional selectivity, lipases are specific to the position of the ester bond in TAG. In this regard, two conditions can be pictured as given in Figure 1.4. Some lipases can catalyze reactions of all hydroxyl groups and thus can release fatty acid at any of the ester bonds in TAGs. These are referred as non-specific lipases and they can catalyze complete breakdown of TAGs to glycerol and free fatty acids. Examples to this kind of lipases are found in *Burkholderia cepacia, Burkholderia glumae, Chromobactierum viscosum, Pseudomanos fluorescens* [17, 18]. 1,3-specific lipases, on the other hand, catalyze reactions of only primary hydroxyl groups of TAG (Figure 1.4). These lipases produce free fatty acids, diacylglycerols and/or monoacylglycerols from 1- or 3- positions of TAG. Most of the microbial lipases are 1,3-specific and examples to the lipase origin include *Bacillus thermocatenulatus, Aspergillus niger, Candidada antartica* B, *Rhizopus orzyae* and *Mucor miehei* [17–19]. In addition some lipase are found to be specific for certain fatty acids and catalyze the esters formed from them. For instance the hydrolysis of only long-chain fatty acids with a double bond between C9 and C10 is performed by the lipases from *Geotrichium candidum* and ungerminated oat seeds. These lipases can be also referred as fatty acid specific lipases [19, 20].

1.4.2 Stereo-selectivity

Stereoselectivity refers to the selective formation of one stereoisomer (enantiomer or diastereomer) over another. Lipases may show insignificant stereo-specificity or they can

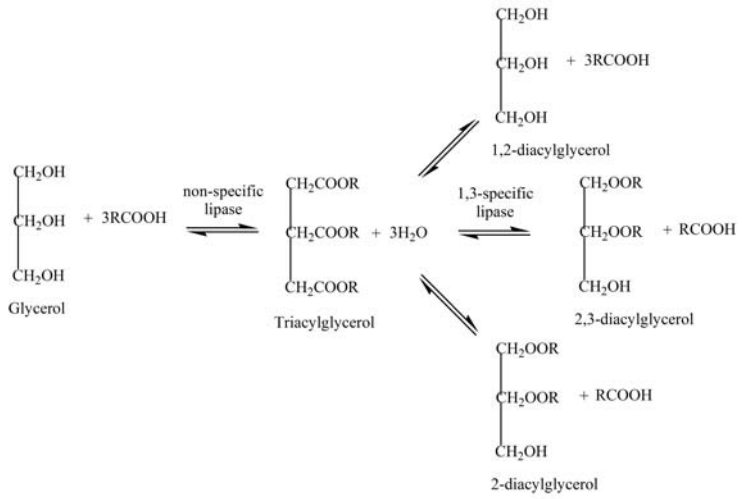

Figure 1.4:
Regio-selectivity in TAG Hydrolysis. Non-specific lipases hydrolyze TAG completely to produce free fatty acids and glycerol. 1,3-specific lipases catalyze the *sn*-1 and/or *sn*-3 ester linkages to a partial hydrolysis of triacylglycerols to free fatty acid, diacylglycerols and monoacylglycerols.

be specific to either *sn*-1 or *sn*-3 hydroxyl groups of triacylglycerols [2]. Stereoselectivity differ from lipase to lipase and from substrate to substrate for single lipase. Two lipases from *Pseudomonas* sp. and *Pseudomonas aeruginosa* show high *sn*-1 preference if the substrate is trioctanoin while most of the other microbial lipases exhibit moderate stereo-selectivity towards trioctanoin. *Candida antartica* lipase B, on the other hand, has a high stereo-selectivity towards *sn*-1 of trioctanoin but it changes to *sn*-3 preference when triolein as substrate. In addition they can discriminate between R- or S- enantiomers of various esters. Both enantio- and regio-selectivities of lipases make them particular attention for selective preparation of enantio-pure chemicals as well as selectively functionalized or protected molecules [21, 22].

1.4.3 Substrate-selectivity

Lipases exhibit selectivity with respect to the type and the chain-length of the acyl groups in their substrates. They can discriminate against certain fatty acids or groups of fatty acids.

For example, porcine lipase is specific to *cis*c-2 over *cis*-7 octadecenoyl moiety [2]. Moreover, the lipase activity towards alcohols is shown to be in order as their most preferred substrate is primary alcohols and the least preferred is tertiary alcohols [23]. Apart from triacylglycerols and aliphatic esters, lipases can also accommodate cyclic esters, thioesters and amines [18, 24, 25]. Additionally, the chain-length selectivity of lipases has been considered in several studies [26–29] and with respect to the chain-length of fatty acids, lipases mostly prefer a range of medium (C6) to long (C16) chain-length [2]. However, few exceptions are encountered to this generalization and these include *Penicillium roquefortti* and *Bacillus thermocatenulatus* lipases which hydrolyze only short chains of C4.

1.5 Structure

The first lipase structure has been crystallized in 1990 by Brady [14]. There are more than one hundred of 3D lipase structures in protein data bank[1]. Based upon these studies, some common structural features are found for all lipases. These can be listed as:

- All of the lipases are a member of α-β hydrolase fold such that lipase structures are composed of central β sheets and surrounding α helices [30–33].

- The catalytic serine in all of the lipases are found in a hairpin turn between an α-helix and α-helix or β-sheet. This region also contains a highly conserved penta-peptide sequence of G-X-S-X-G which forms a characteristic turn that is also referred as ”nucleophilic elbow” [31, 32, 34].

- The active site in lipases is common to another class of hydrolases; serine-protease which also consists of three amino acids, serine, histidine and aspartic/glutamic acid [14, 35]. Although proteases and lipases share the same chemistry for active

[1] According to release statistics of PDB in 2012.

site, the structural arrangement of the residues in active site is oriented to invert the stereochemistry of the catalytic triad in lipases compared to proteases [36].

- An amphiphilic lid that covers the active site is found in lipases [37, 38]. The composition and size of the lid structure vary among other lipases. For instance, the lid from the lipase of guinea pig has only five amino acids while the lid from *Bacillus thermocatenulatus* lipase consists of two α-helices which corresponds to 20% of the whole lipase structure [33, 39].

- The catalytic cleft of lipases is found to composed of four binding pockets, three pockets holding *sn*-1, *sn*-2 and *sn*-3 acyl chains of triacylglycerol and an oxyanion hole which is formed by the constellation of the amine groups to stabilize the tetrahedral anion during catalysis [40, 41].

Apart from these characteristics of lipase structures, the catalytic cleft in lipases are attributed to specificity. Pleiss et al. investigated the lipase clefts, particularly the fatty acid binding sites, and found that lipase structures may exhibit three different geometry; crevice-like, funnel-like and tunnel-like [41]. These differences in the binding pockets account for the diverse substrate specificities in lipases. In addition stereoselectivity in lipases is dependent on the steric interaction of the cleft with the substrate and thus lipase structure is also critical to understand the stereoselectivity [42].

1.6 Interfacial Activation

In 1958, Sarda and Desnuelle investigated the lipase activity and hypothesized ”interfacial activation” of the lipase in the presence of water insoluble substrates through a conformational change in the lipase structure [43]. The direct evidence on the interfacial activation was obtained from the crystallographic investigation of *Rhizomucor miehei* lipase complexes with two different ligands [37, 44]. These lipase crystals that were captured in the

activated form, are accepted as proofs for a conformational change in the surface loop with respect closed structures revealing the structural basis of the interfacial activation. Following this finding, a list of other lipases were crystallized and reported similar rearrangements in the activated lipase structure. In the following years, many biochemical and biophysical methods had been developed to study this phenomena and they have provided a wealth of information on the macroscopic properties of lipase activity [38, 45–47].

If we consider the classical Michaelis-Menten kinetics which relies implicitly on the assumption that enzyme and substrate are in the same phase i.e. both are soluble, one cannot apply this model directly to the lipase activity at the water-lipid interface. Instead, the lipase activity at the interface is an example to heterogeneous catalysis which is dependent on many factors such as interfacial organization, steric conditions of the lipid interface and lipase structure and the physical interaction between the reactant molecules. In principle, the chemical interactions of the lipase and the substrate is coupled with many other physical processes such as adsorption, desorption and diffusion of the reactants and products [47]. Thus, the reaction rates may be affected by one or any combination of such physical properties. To date, two mathematical schemes are proposed for the lipase kinetics at the interface (Figure 1.5). According to the simplest model proposed by Verger et al. (1976), soluble lipase (E) is adsorbed on the insoluble substrate interface in a reversible manner and fixed at the interface [45]. This adsorption leads to a more favorable energy state for enzyme (E*) and after this, the kinetics fits into simple Michaelis-Menten scheme. The enzyme-substrate (E*S) complex is then formed and followed by the formation of products and their instantaneous solubilization at the interface or in aqueous media.

According to the second model proposed by Martinelle and Hult [48], the lipase exists in two conformations i.e. open and closed in solution. These two states are in equilibrium favoring the closed form in solution and when an interface is encountered the equilibrium changes towards the open form via the adsorption of the lipase to the interface which is

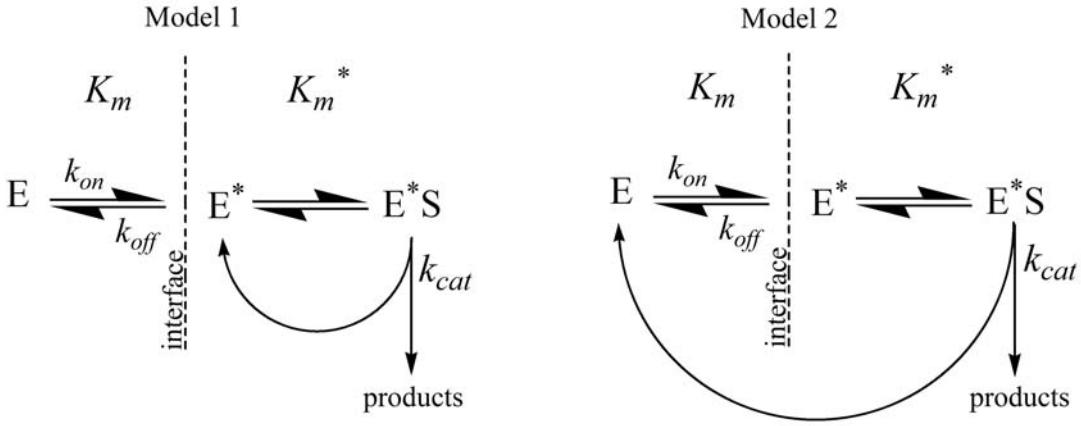

Figure 1.5:
Models for Interfacial Kinetics. (Model 1) Enzyme stays in the interface during catalysis or (Model 2) returns to bulk water by dissociating from the interface after product formation. The dashed line is for two dimensional representation of the interface. For both models, the products are released to bulk water or stays in the interface with respect to their physical properties. This scheme is adapted from ref. [50].

stabilized by the hydrophobic environment of the interface.

Understanding the kinetics of TAG hydrolysis by lipases is essential for rational design purposes to modify catalytic characteristics. However, it may be experimentally labor-intensive since hydrolysis of water insoluble TAGs takes place at the water-lipid interface and these reactions follow interfacial kinetics rather than simple Michealis-Menten model [45, 49]. According to the proposed models (Figure 1.5), the lipase is activated by adsorption to the substrate interface which is composed of several TAG molecules [50]. Then a single TAG binds to the catalytic cleft to form enzyme-substrate (ES) complex of hydrolysis and this binding is critical for assessing catalytic preference ($\Delta G_T^{\ddagger}$) of lipase action at the interface [50]. Unfortunately experimental quantification of single TAG binding is quite labor-intensive due to the physical properties of lipid substrate [32, 50].

1.7 Lipases in Industry

Today we know that enzymes are completely biodegradable and renewable as they are typically produced by fermentation from bio-based materials [51]. Additionally, recent advancements in the production processes in the last decade show that it is possible to obtain enzyme concentrates at relatively low costs [52] and these efforts have opened new doors for enzyme to be applied in bulk industrial process.

Naturally occurring fats and oils are the preferred substrates of lipases yet they can still catalyze hydrolysis or production of wide range esters [53–56]. Owing to their wide spectrum of substrates, they are one of the largest class of biocatalysts since two of the dominating areas of enzyme market -with respect to volume and value of the enzyme-, food processing (40-45%) and detergent industry (35-40%) are mainly the performance of lipases [55]. Additionally, novel biotechnological applications has been established using lipases for synthesis of biodiesel, polymers, drug intermediates, agrochemicals, amino acid derivatives and flavor compounds [55, 57], as well as in biosensors and bioremediation processes [58–60]. Table 1.1 explains the most promising lipase applications in various industrial sectors.

Lipases can catalyze reaction in aqueous and also in organic media. This has made them particularly attractive for solving challenging synthesis of organic reactions for the last decade [61, 62]. They have a widespread use in industry (Table 1.1) consequent to their commercially important reactions/products and this makes them particularly attractive for protein engineering applications. Because of their extracellular nature, microbial lipases are easy to produce at large quantities and to isolate. Particularly, microbial lipases are the preferred source for many industrial applications. The major reasons for the high industrial potential of microbial lipases include, (i) they are relatively more stable in organic solvents, at high temperatures and ionic strengths, (ii) they do not require cofactors (iii) they have a broad substrate selectivity and high enantioselectivity [59]. The interest

in microbial lipases is reflected by an large number of articles, reviews and monographs that studied in the molecular biology, biochemical properties and structure of lipases along with their biotechnological applications [14,18,48,50,63–73]. Examples of such commercially microbial lipases are listed in Table 1.2. Lipases are used in many industrial sectors that as in production of cheese flavor as in food industry, of detergents as in oleochemistry and synthesis of various esters as in organic synthesis. The lipase from *Penicillium roquefortti* which has high activity towards tributyrin, the most abundant fat in milk, is used for production of the flavor in the *Roquefort* cheese [74].

Apart from the listed examples in Table 1.2, many other known applications of lipases include removal of the pitch in pulp industry. Pitch is defined as the hydrophobic component of wood and must be removed before processing. Enzymatic removal of pitch is performed via *C. rugosa* lipase at low costs [75–77]. Textile industry also uses lipases to bypass the process called ”stoning”, particularly for denim production [78]. The advantages of this enzymatic bypass are numerous including reduced exposure to toxic chemicals used in chemical process which requires asbestos. Lipases are also used in pharmaceutical industry due to their ability to synthesize of enantio-pure drugs [79–82]. Examples to this kind of drugs include Ibuprofen for pain, Taxol for cancer and Diltiazem for high blood pressure medications. Enantioselective production of drugs by lipases reduces costs and drug toxicity in pharmaceutical industry.

Table 1.1: Industrial Applications of Lipases.

Lipase Reactions	**Industry**	**Product/Usage**	**References**
Hydrolysis of fats	Detergent	Stain removal	[83]
Hydrolysis	Dairy, paper and leather	Flavor, paper and leather quality	[48, 84]
Transesterification fats/oils	Natural oils	Cocoa butter	[85]
Glycerolysis of fats/oils	Surfactant	Monoacylglycerols	[18]
Esterification/Transesterification	Flavor, fragrance	Synthesis of natural flavor ester	[86]
Acylation of sugar alcohols	Surfactant	Sugar monoacyl esters	[87]
Transesterification	Fuel	Biodiesel	[88–90]
Racemic resolution	Pharmaceutical	Chiral drugs, agrochemicals	[83, 91, 92]

Table 1.2: Commercially Important Microbial Lipases.

Origin	Organism Name	MW kDa	Features	Industry [18, 55]	Commercial Source [13, 18, 73]
Fungal	*Candida rugosa*	60	nonspecific	Organic synthesis	Amano, Biocatalysts, Boehringer, Fluka, Mannheim, Genzyme, Sigma
	Candida antartica A/B	60	*sn*-1,3	Organic synthesis	Boehringer Mannheim, Novo Nordisk
	Rhizomucor miehei	30	*sn*-1,3	Food processing	Novo Nordisk, Amano Biocatalysts
	Humicola lanuginosa	30	nonspecific	Detergent additive	Novo Nordisk, Mannheim, Boehringer
	Fusarium solani	22		Oleochemistry	
Bacterial	*Pseudomonas mendocina*	33	*sn*-1,3	Organic synthesis	Genencor
	Pseudomonas pseudoalcaligenes	33	*sn*-1,3	Oleochemistry	Genencor
	Pseudomonas glumae	33	nonspecific	Organic synthesis Oleochemistry	
	Burkholderia cepacia	33	nonspecific	Organic synthesis	Amano, Fluka, Boehringer Mannheim
	Bacillus thermocatenulatus	43	*sn*-1,3 thermophilic		

1.8 Why Engineer Lipases?

Enzyme catalysis inevitably obeys the thermodynamic rules set for chemical catalysis but what makes enzymes special is basically their efficiency and selectivity. Using enzymes as catalysts, the synthesis of specific products is possible due to the selective nature of enzyme reactions. Selectivity of industrial production greatly reduces waste and energy consumption, resulting in a much greener manufacture. In this sense, enzyme catalysis is superior to chemical catalysis and as a consequence enzymes are important for many industrial applications.

With numerous applications in range of fields (Table 1.1) in bulk industrial processes (Table 1.2), lipases can be engineered to produce new enzymes with altered substrate selectivities and improved stabilities. Such lipase variants may substantially improve efficiency of the industrial biocatalysis. Hence, lipases are one of the most popular class of enzymes that is considered in protein engineering studies [92]. Still, the use of lipases in industrial processes has some limitations. For instance, harsh conditions faced in industrial processes can be detrimental to the protein nature of lipases. To overcome this particular limitation, thermostable lipases are current interest of lipase research [92, 93].

2 PROTEIN ENGINEERING

2.1 Background

Protein engineering is the research field that concerns design of new proteins and enzymes with improved properties and has a unique place for today's science and technology. The engineering approaches may rely on different principles, trading off rational approach using three-dimensional (3D) models against random approach using directed evolution [94]. Many fundamental questions can be answered via protein engineering such as protein folding/stability and structure/function complementarity [95–97]. Along with its research and medical impacts, protein engineering has wide variety of industrial applications. Important bioproducts such as vaccines, replacement therapeutics and industrial enzymes are improved and produced via protein engineering. Essentially, the significant growth -50% per year- in the U.S. enzyme market between 1992 and 1997 was a result of successful industrial applications of enzymes such as proteases, lipases, substillins and cellulases [94].

2.2 Setting up Rationale for Engineering

A successful application of protein engineering requires knowledge on the composition and production of proteins and is usually performed in three steps which are design, mutagenesis and production. Before altering the composition of the wild-type enzyme, it is essential to have a well-defined rationale for the design. In literature, there are numerous examples utilizing wide range of methods for enzyme engineering purposes however, the engineering rationale may be based either on the selection of the improved enzymes from pools of randomly generated variants or on the rational design of new enzymes with improved

properties [94]. Table 2.1 lists widely used tools for random and rational approaches. The random approach is generally accompanied by random mutagenesis which is the laboratory practice of evolution for a DNA fragments and thus often referred as directed evolution. The method is based upon generation of a library of mutants and its success is principally dependent on the screening method. As explained by F.H. Arnold, a convenient and rapid screening strategy for variations is required for random approach i.e. ”you get what you screen for” [98]. In random approaches, the design and mutagenesis are performed at the same time and rely on directed evolution which is multiple random mutations introduced successively into a DNA piece [99], or domain-swapping which is shuffling of multiple genes [100], or circular permutation which is shuffling of protein termini (2.1) [101].

In rational approach, design and mutagenesis steps are separate. In the design, the rationale is established mostly via computational analysis of three-dimensional structures i.e. molecular docking and molecular dynamics approaches. The mutagenesis approaches in rational design is usually site-directed. The selected sites may be substituted with other amino acids (site-directed mutagenesis) [102] or saturated with the rest of other 19 amino acids (site-saturation mutagenesis) [103] The rational approach is based on an educated guess made by utilizing high resolution information on the enzyme of interest [104].

Choosing between random and rational approaches rely on the amount of the starting information [98]. Random design approaches utilizing evolution mechanism are applicable to cases where only little information, i.e. primary sequence is known. If the resolution of the starting information is high, for example the crystal structure is available, the popular choice is rational design. The outcomes of the both approaches also differ such that it is possible to comment on the structure and function relation of the enzymes in rational design whereas the output is dependent on the screening in random design.

After the design step i.e. selecting critical regions in the enzyme in a random or rational approach, the straightforward step, mutagenesis is followed for both approaches. The final

Table 2.1: Protein Engineering Approaches.

	Random	**Rational**
Design	Directed evolution Domain swapping Circular permutation	3D models Molecular docking Molecular dynamics
Mutagenesis	Random Gene-shuffling	Site-directed Site-saturation
Production	Heterologous protein expression Purification	

step is the production of improved enzymes and their relevant characterizations to confirm improvements. The production stages of both random and rational approaches employ similar methods which are heterologous expression and/or purification of the recombinant protein.

2.3 Rational Approach

In rational approach, the first step; design employs mostly computational methods for prediction of the regulatory regions in enzymes, which may be important for the stability, function, selectivity of enzymes. After design, site-directed mutagenesis and heterelogous protein expression techniques are carried out to obtain the enzyme analogues. Next, in this chapter, a methodological overview for the steps of the rational approach are given as design (subsection 2.3.1), mutagenesis (2.3.2), production (2.3.3).

2.3.1 Design

In rational approach, design of novel enzymes is the first step which involves selection of critical sites responsible for a given character of the enzyme. The success of design is limited to the resolution of information on the enzyme of interest. This approach works best if the 3D structure of protein or enzyme is available. The structures in which the enzyme is captured during an act, i.e. bound to its substrate analog, is particularly beneficial for

predicting critical regions for enzyme function. Molecular docking approaches may assist in the prediction of such regions in the enzyme surface that rules molecular recognition i.e. substrate/inhibitor binding. In addition, significant progress in molecular dynamics approaches enables us to see 3D models for enzyme action. Thereof such computational approaches are significant tools for enzyme engineers during design. Design is an important test of our understanding of the enzyme structure and function. Usually, in the design procedure, there is an obvious logic: choosing sites are close to the active site, which fit with the proposed reaction mechanism, sites in the binding pocket for the substrate, or sites that make important structural contributions to the enzyme for example stability. To select such sites, one need to examine the enzyme structure thoroughly. Thus, the availability of the 3D structure of the enzyme is essential in this approach. Currently, the Brookhaven Protein Data Bank [105] is built with great deal of structural information on diverse proteins including enzymes[1]. After determination of the 3D structures, computer assisted approaches are now an indispensable tool for protein studies since it provides novel insights to complex systems which may be inaccessible to experimentation.

Technically speaking, proteins are sets of Cartesian coordinates varying at positions (x, y, z) of N, O, C, S and H atoms of hundreds of amino acids that constitute them. Even if a small portion of protein structure is considered for analysis, computers are the only convenient tool to deal with such large number of datasets. Computational methods provide powerful service for understanding the atomistic details of biological phenomena ruled by proteins and their partners. Computers can also be used as a convenient viewing interface examining the very details of protein structure. The advances in the computer assisted methods have a direct impact on the structural analysis of enzymes [106–108]. Molecular docking approaches could be used to model enzyme-substrate interactions whereas other methods developed for estimating the free energy changes from multiple steered molecular

[1]Enzymes compose 20% of all PDB entries according to PDB statistics released in 2012.

dynamics trajectories [109]. Following subsections, (2.3.1, 2.3.1, 2.3.1) gives an overview of three of the computational methods used in this thesis. Few methods related to design are explained in the subsections.

Molecular Docking

In molecular modeling, docking is a way of predicting the orientation of a molecule with respect to a second when they are bound to form a complex. Using scoring functions, knowledge of the preferred orientation can be used to determine the strength of binding i.e. affinity between two molecules. Docking plays an important role in the rational design of inhibitors for receptors as well as substrates for enzymes [110, 111]. Considering the biological and pharmaceutical significance of molecular docking, considerable efforts have been dedicated to improve the docking methods. Two main categories for molecular docking can be listed as: (i) rigid-flexible docking which allows only ligand (substrate or drug molecule) to change its orientation (depending on its torsional degree of freedom) during docking calculation and (ii) flexible-flexible docking which also allows macromolecule (protein or enzyme) to change its orientation for example around the active site.

A number of algorithms are developed for molecular docking purposes, these include AutoDock [112], FLEX [113] for small ligands and HADDOCK [114] particularly developed for protein-protein docking. AutoDock, the most cited small molecule docking algorithm, predicts dominant intermolecular complexes of proteins and ligands of known 3D structure. Currently the software is distributed in two versions AutoDock 4 [112] and AutoDock Vina [115]. AutoDock Vina is better suited to large ligand docking with increased accuracies at improved speeds compared to AutoDock 4. In this method, the ligand-protein interactions by means of electrostatic interactions, hydrogen bonding and desolvation energies are calculated over a pre-defined grid which includes the binding region for ligand [116]. The Genetic algorithm is used to search for the global minimum of

the bound complex.

Molecular docking approaches are used to estimate the 3D model for the complex molecules in biological systems and spotting such complexes holds importance by means of their biological relevancy as displayed previously in examples [117–119].

Molecular Dynamics

Molecular Dynamics (MD) is a powerful computational method that describes the equilibrium and dynamics of a biological system via generating configurations of that system by integration of classical mechanics to calculate the time dependence of the system. MD generates information at the microscopic level which includes atomic coordinates and velocities and the simulations are used to understand the dynamics of the biological system. In this respect, MD simulations can be used to resolve structure-function relationship of proteins.

MD technique is widely used in the context of molecular biology. Protein dynamics is first introduced by McCammon et. al. in 1977 through an analysis using bovine pancreatic trypsin inhibitor [120]. Now MD has widespread use for several purposes including refinement of X-ray and NMR structures, searching for new conformations and calculations of binding energy.

Building and/or manipulating a protein structure in the computer requires knowledge of the bond, angle, torsion and atomic radii of the protein structure. MD simulations are performed with a set of force fields that corresponds to the potential energy of the static protein structure. The potential energy of a rigid protein was previously described by a number of force fields such as ECEPP [121], Hagler's [122, 123], Herman's [124], Allinger's [125], MM2 and CHARMM [126], AMBER [127], GROMOS [128] force fields. These force fields are employed to built a wire model for a protein and the potential energy of the protein molecule is obtained by a number formulations that computes bonded i.e. rota-

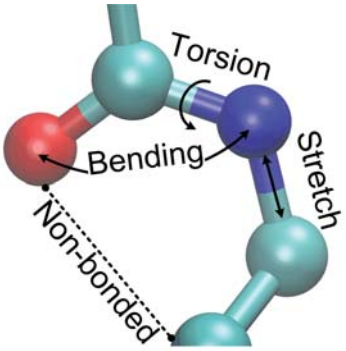

Figure 2.1:
Energy terms for Simple Molecular Mechanics. Non-bonded atoms (greater than two bonds apart) interact through van der Waals attraction, steric repulsion, and electrostatic attraction/repulsion. Bonded interactions in a simple mechanical force field include stretching, bending and torsions.

tions and torsions, and non-bonded i.e. van der Waals and electrostatic potentials (Figure 2.1). For the time dependent change in the bond lengths, bond angles and torsions, also the non-bonding van der Waals and elecrostatic interactions between atoms, force fields, a collection of formulations are used.

At any temperature above absolute zero, the total energy of a protein molecule is composed of potential energy and kinetic energy of the thermal motions of the atoms. This motion can be generated via computers by calculating the derivatives of the potential energies i.e. velocities, accelerations and forces at time $t_{+}\Delta t$ by integration as in classical mechanics. Smooth motion trajectories are obtained for very short integration steps, mostly on the order of femtoseconds (s^{-15}) and a molecular dynamics trajectory is the collection of the successive frames of the atomic positions and velocities. In molecular dynamics, these empirical force fields can also be coupled with a random drag force which may be user-defined to study unfolding and binding of ligands. Development of force fields as mathematical expressions for these features which is the field of molecular mechanics, result in many software packages for MD simulations of biomolecules [129]. NAMD, the most widely used software package, is used in this thesis [130].

In MD simulations, the molecules are surrounded by boundaries which act as symmetry

operators to generate periodic boundary conditions. These periodic boundary conditions enables the simulation to be completed using a relatively small number of particles such that these particles experience forces as they are in bulk solvent.

MD has broadened our understanding of motions of biological molecules which are often inaccessible for experimentation. It permits the study of complex and dynamic processes in biological systems such as folding, stability, conformational changes and transport mechanisms. Although, longer simulation times and larger systems requires more computational power, speed of computers and parallel computing algorithms are getting to more advanced levels improving the accuracies of computer modeling techniques.

Nonequilibrium Methods for Free Energy Calculations

Most biological complexes shuttle between two equilibrium states at time scales that are beyond the reach of current MD simulations, their equilibrium dynamics cannot be replicated in a straightforward manner. Nevertheless one can overcome this problem by using biasing or adaptive methods. Here, the steered molecular dynamics (SMD) together with Jarzynskis equality [109] which is a popular non-equilibrium method is explained [131].

In 1997, Jarzynski has postulated an equality between nonequilibrium work functions and equilibrium free energy differences in a microscopic system, which provides a powerful tool for the determination of equilibrium free energy changes in the context of steered molecular dynamics simulations. This equality has so far intensely applied for several objectives including ligand binding to protein [109, 132–141].

The SMD method is based on guiding the system from one equilibrium state to another by a drag force [142, 143] which can be seen as an *in-silico* counterpart of single-molecule experiments that utilize atomic force microscopy or laser tweezers [144, 145]. According to Jarzynskis equality, the Boltzmann average of the work functions obtained from the SMD simulations provides an estimate of the potential of mean force of the ligand along

the reaction coordinate, which approaches the true value when the system is sampled sufficiently [109]. During SMD simulation, one part of the system, i.e. ligand, is coupled with a harmonic oscillator. The spring is pulled to build up a certain tension on the ligand. The movement of this spring builds up a certain tension which results in a molecular transition i.e. unbinding and unfolding [146]. Because SMD is a path-dependent method, it can also be used in searching for a ligands pathways [136, 147]. Furthermore, the trajectory data obtained from SMD simulations provide a wealth of information which can be utilized, for example, in finding the critical enzyme residues that strongly interact with a ligand along its path. Such information would be very useful in rational design of enzyme analogues with improved selectivity properties for a particular ligand.

While SMD simulations are less laborious and easier to implement compared to other path-dependent approaches such as umbrella sampling MD simulations, they require relatively more computing time to obtain the same level of accuracy [148, 149]. Nevertheless, this approach may be applied for the cases where the comparison of two systems with either different ligands or different macromolecules is considered. Additionally, inaccuracies in force fields such as missing polarization interaction [150, 151] are expected to affect the absolute binding free energies of ligands but not their relative values. The nonequilibrium methods like Jarzynskis equality are obviously useful for the cases where nonequilibrium conditions are dominating and cannot be easily avoided. Following the Jarzynski's equality, other approaches are also developed for similar purposes [152].

2.3.2 Mutagenesis

Mutagenesis is one of the most widely used DNA manipulation techniques. Via mutagenesis, one can investigate the impact of sequence alterations in the DNA and also in the downstream process of DNA expression i.e. mRNA, protein. Today, it is possible to select from a wide range of mutagenesis approaches that include chemical, oligonucleotide-,

polymerase chain reaction (PCR)-based and cassette mutagenesis [94]. Among these mutagenesis ways, site-directed mutagenesis which is widely used for protein engineering purposes is focused here.

Site-directed mutagenesis is an in-vitro method based on PCR. It uses PCR, more specifically, overlap extension PCR (OE-PCR) to generate specific mutations at predetermined locations of a DNA piece [102]. Figure 2.2 illustrates the principle of site-directed mutagenesis via OE-PCR. The mutant sequences are carried in the primers which are used for amplification of mutant DNA. The second round is another PCR uses first reaction products as primers and template. For about 10-20 steps -depending on the length of the overlap region- the second PCR does not require primers but after some point, the efficiency of amplification can be enhanced using gene-specific primers annealing at the 5' and 3' of DNA piece. Today, there are also other kit based approaches which are relatively easier and faster because these kit-based approaches bypass cloning steps [153]. Lastly, artificial gene synthesis is now relatively cheaper as it can be provided by numerous companies worldwide. Gene synthesis may save time and money compared to mutagenesis procedures. But for a large number of mutants, these kit based approaches and gene-synthesis may be costly.

2.3.3 Production

After mutagenesis, next aim is to produce recombinant proteins (wild-type and/or mutants). Heterologous protein expression is used for this purpose owing to the availability of various expression systems such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Pichia pastoris*. *E. coli* will be given special emphasis here and in the thesis.

The host, *E. coli* because of being a prokaryotic species has resulted in loss of about 70% of eukaryotic protein expression but this host is still the first choice in most of the protein engineering approaches [94]. Since, *E. coli* is actually the organism on which the recombinant DNA technology is founded and there are lots information on the mechanism

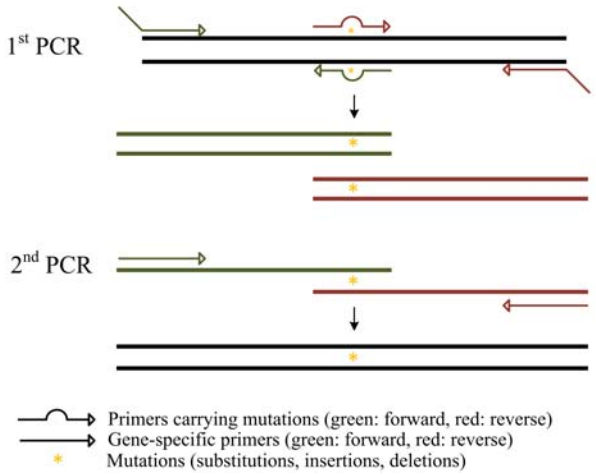

Figure 2.2:
Principle of OE-PCR. A graphical illustration of OE-PCR method and the corresponding templates and primers for sequential PCRs are shown. The legend is given at the bottom of the figure.

and machinery of the *E. coli* expression system. In this regard, *E. coli* may be an ideal first choice for an expression study. Moreover, co-expression studies showed that *E. coli* can express more than one protein at the same time [154]. The choice of expression host is also dependent on the enzyme of interest. For eukaryotic proteins, mostly yeast systems are preferred due to their ability to perform post-translational modifications on the recombinant proteins such as phosphorylation and glycosylation. Other than this, there are several factors need to be looked in expression study such that the choice of promoter and strain may directly affect the efficiency of expression. Moreover, in order to obtain soluble forms of the recombinant enzyme product, the gene sequence for it may be fused with another gene piece which produces soluble proteins such as glutathione-S-transferase and maltose binding protein.

The production may be scaled up according to the needs: a cellular protein can be expressed to study its function in cell biology, bench-scale production of protein can be achieved for characterization studies or the production can be scaled up to fermentation for industrial processes.

After design and mutagenesis, last step is to obtain the recombinant products which is followed by the characterization of the mutations. A range of methods are available to obtain information from the macroscopic level i.e. protein size, molecular mass, enzyme activity.

3 THESIS APPROACH

3.1 Purpose

Schmit-Dannert et al. (1997) investigated two different lipases originating from the bacteria, *Bacillus thermocatenulatus* (BTL1 and BTL2) and they showed that these lipases have distinct features from each other [155]. BTL2 stood out with its higher activity and stability compared to BTL1 and other known lipases. BTL2 is of particular interest for this thesis. Here we report the results of the two case studies on the protein engineering of BTL2. Chapter 5 aims for the rational design of BTL2 to increase its selectivity towards C4 TAG and Chapter **??** aims for the investigation of the impact of two tryptophans in the lid region of BTL2 in terms of activity, thermostability and aggregation. For both chapters, computational methods are implemented using two atomistic models of BTL2 (PDBID: 2W22 [33] and 1KU0 [156]. The former is used in Chapter 5 and the latter is in Chapter **??** (Figure 3.1).

3.2 Thesis Approach

Lipases, an enzyme class, is explained in the context of their reactions, specificities, structures and industrial importance in the first chapter. An overview of enzyme-engineering is given in the second chapter. Here, in this chapter, thesis approach and thesis aims are explained. The following chapters are related to two protein engineering studies concerning the rational design of BTL2 to enhance the catalytic preference and the semi-rational design to investigate the function of lid tryptophans in BTL2. Chapter 5 aiming to increase catalytic preference of tributyrin (C4) hydrolysis of BTL2 involves molecular docking ap-

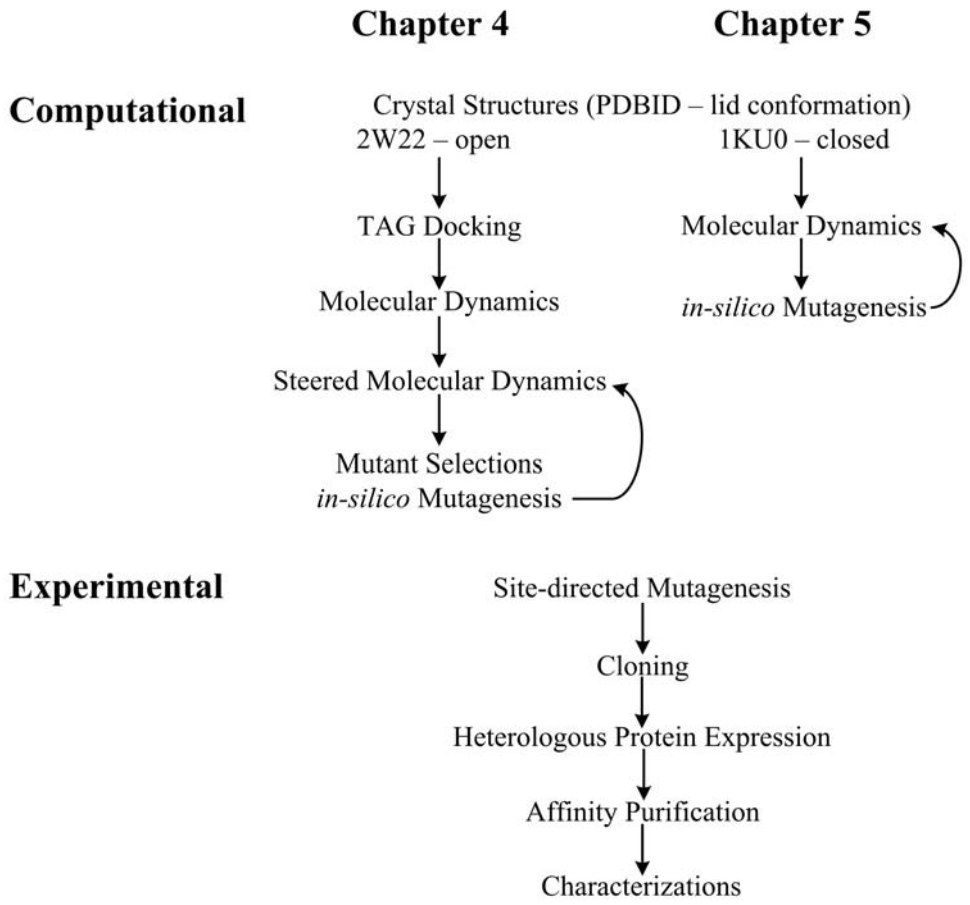

Figure 3.1: Thesis Methodologies

proach to determine the bound configuration of two different TAG molecules at length of C4 and C8 to the catalytic cleft of BTL2 and molecular dynamics (MD) of these complexes. Following this, steered molecular dynamics (SMD) is performed to model the dissociation of the bound ligands from the catalytic cleft (Figure 3.1). Upon identification of ligands' path, selected residues which are confined to the catalytic cleft are tested *in-silico* using the same SMD approach and *in-vivo* using the molecular biology methods for cloning, site-directed mutagenesis and heterelogous protein expression (Figure 3.1)to modify the catalytic efficiency of lipase. The chain-length selectivity is explained in terms of the change activation barrier of mutants with respect to TAG chain-length to finalize the rational design. Chapter **??** is related to molecular investigation of the lid tryptophans in BTL2. The lid region in lipases are rich in aromatic residues to mediate the hydrophobic interactions with the environment. Thermophilic enzymes like BTL2 stay stable at high temperatures by intra-molecular hydrophobic interactions [157]. BTL2 which is originated from a thermophilic microogranism could be used to understand the basis of the thermostability provided by the hydrophobic interactions. These interactions might be critical to the activity, stability and aggregation tendency of lipases [158, 159]. BTL2 has a large portion of hydrophobic residues, especially it contains two tryptophans (W212, W235) in the lid domain. Chapter **??** investigates the effects of these tryptophans on the thermostability and aggregation of BTL2 using similar protein engineering tools (Figure 3.1). For both chapters, computational methods are implemented using two atomistic models of BTL2 (PDBID: 2W22 and 1KU0). The former is used in Chapter 4 and the latter is in Chapter 5 (Figure 3.1).

3.3 *Bacillus thermocatenulatus* lipase

Phylogenetic, biochemical and structural information on BTL2 is collected from the relevant literature and given next. In addition, experimental studies are performed with wild-

type BTL2 ahead of protein engineering to find out the optimum expression and purification system.

3.3.1 Lipase Family 1.5

According to the classification of bacterial lipases [160], eight lipase families are identified with respect to their peptapeptide motif containing catalytic serine. Five of the six subfamilies found in Family 1 has G-X-S-X-G pentamer while members of Lipase 1.5 has conserved the A-X-S-X-G motif. Although some features of this subfamily lipases are divergent from other subfamilies in Family 1, the members of Family 1.5 have significant sequence similarity and Figure 3.2 shows the multiple alignment of previously identified lipases from Family 1.5. With a high sequence homology, these lipases also share common biochemical and structural features. As these lipases are mostly originated from a thermoalkalophilic species, they work best at high pH and high temperatures [161, 162] Moreover, this family lipases are larger lipases (more than 40 kDa) [160] compared to other subfamilies because they possess an extra domain for zinc coordination which accounts for increased thermostability and also not found in other lipase families [156]. Owing to some unique and conserved characteristics, members of family 1.5 have been a particular interest of enzyme engineers [163, 164].

3.3.2 Heterologous Expression Trials

Before engineering of this lipase, the native BTL2 is expressed in two heterologous hosts: *E. coli* and *P. pastoris*. Among these two hosts, *E. coli* is preferred for BTL2 expression -although both of the expressions are successful- since *E. coli* system is the most convenient way of heterologous expression. In these *E. coli* expressions, fusion tags are used (Figure 3.3). These include poly histidine (six-His), glutathione-S-transferase (GST) and maltose binding protein (MBP). Such tags are important for the solubility of recombinant

Figure 3.2:
Multiple Sequence Alignment of Lipase 1.5 Family. The penta peptide motif that holds catalytic serine is shown in lines and catalytic triad are also shown with dots. The secondary features are also shown on top of the alignment. The sequences of two *Bacillus stearothermophilus* lipases (PDB ID: 1JI3, 1KU0), *Bacillus thermocatenulatus* lipase (PDB ID: 2W22) and *Geobacillus* species SBS-4S lipase (3AUK) are used in the alignment.

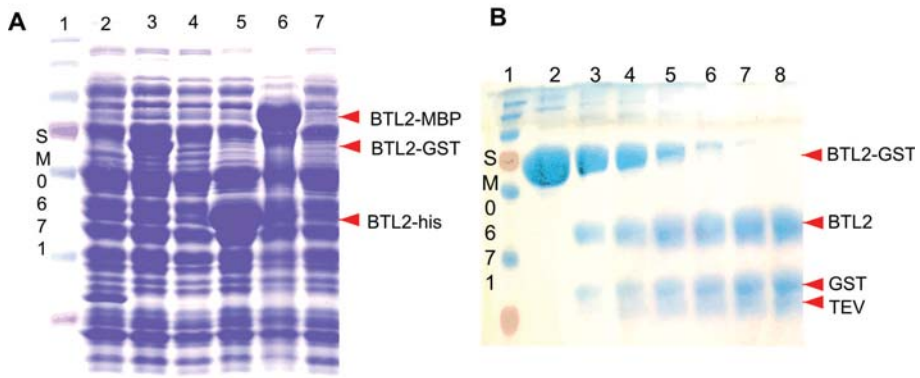

Figure 3.3:
E. coli Expression of BTL2 in Native and Fusion Forms. 12.5% SDS-PAGE analyses are shown for (A) soluble *E. coli* lysate and (B) native BTL2. (A) Lane 1 is the molecular weight marker (Fermentas SM0671, see Appendix A.1) and lane 3, lane 5 and lane 6 corresponds to BTL2 expression in fusion forms of respective tags: GST, MBP and poly histidine. The expected positions for the fusion proteins are marked with red arrows (BTL2: 43 kDa, GST: 30 kDa, MBP: 50 kDa poly histidine: 1 kDa). (B) BTL2-GST fusion is cleaved by TEV protease whose recognition site is located in the C-termini of GST. Lane 1 is same marker. The amount of the protease is increased gradually from lane 2 to 8. The corresponding molecular size for BTL2-GST (73 kDa), TEV (26 kDa), GST (30 kDa) and native BTL2 (43 kDa) is marked. Protease cleavage is performed in 100 mM Tris-EDTA buffer at pH 8.0 with 1 μM DTT.

protein and may be increase the efficiency of expression [165–167]. These tags also ease the purification steps which is essential for protein engineering studies [168]. But in a typical protein engineering study, it is important obtain native lipase rather than its soluble but fusion forms. Native BTL2 is shown in Figure 3.3. It is obtained after proteolytic cleavage of the GST tag by Tobacco Etch Virus (TEV) protease as described in ref. [169]. This native enzyme is compared with its smallest fusion form (BTL2-his) in a set of experiment which are analyzing lipase activity and thermostability. These tested biochemical features do not differ in between native and poly histidine fused BTL2. Considering affinity purification of tagged proteins, lipase expressions are conducted via poly histidine fusion in *E. coli* throughout this thesis.

3.3.3 Biochemistry

An 1167-bp fragment codes for the mature thermoalkalophilic lipase. It has composed of 389 residues and a predicted molecular weight of 43 kDa. Lipase was found to be stable at pH range of 9-10 and temperature range of 60-70^oC [155]. Optimal working pH range for BTL2 is 8-9 and temperature is 60^oC [170] BTL2 preferentially catalyzes *sn*-1/3 acyl chain in triglycerides similar to other microbial lipases. BTL2 has a narrow chain-length selectivity profile that it exhibits activity towards short and medium chains (C4-C8) with a distinct preference to tributyrin (C4) but it has low activity towards long chains of triglycerides (C10 and longer) [162]. BTL2 is also stable in organic solvents and detergents [162]. In a set of experiments, prior to mutagenesis, we have also analyzed some of the biochemical properties of native lipase tagged with poly histidine. Effects of pH on lipase activity is given in Figure 3.4 A. The pH stability analysis is performed at the pH range of 4-10 is provided in a tri-buffer system [171] which allowed us to eliminate the buffer composition effect on stability of lipase and BTL2-his is found to be stable at the pH range of 7-9 in our experiments which is in agreement with the native BTL2. (Figure 3.4 A) The effects of 11 different metals (Al, Ca, Cd, Cu, Fe, K, Li, Mg, Mn, Na, and Zn) in form of chloride salts are investigated in a concentration gradient. (Figure 3.4 B). Ca, Cd, K, Li, Mg and Na metals do not affect the activity of this lipase in the pure or lysate form whereas BTL2-his is observed to lose a quarter of its activity even at the lowest metal supply of 1 mM for Al, Fe, Mn and Zn presence.

3.3.4 Structure

The crystal structure of BTL2 is resolved at 2.5 Å [33] This crystallization study attracted profound attention, due to the fact that there were no open-state structures, meaning the binding pocket for substrate is available, belonging to lipase family 1.5. Similar to most microbial lipases, BTL2 is a member of α-β hydrolase fold family [172] The structure has

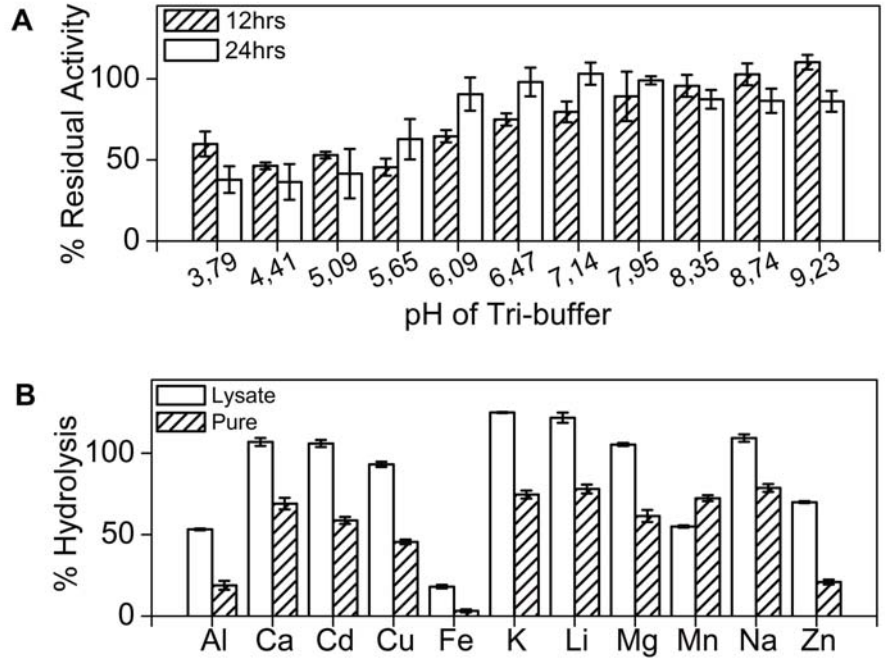

Figure 3.4:
Effects of Metal and pH on Activity.(A) Effect of pH on lipase activity. The pH stability of BTL2 is tested by incubation for 12 hours (patterned bars) and 24 hours (white bars) at room temperature. Residual enzymatic activity is determined in tri-buffer system [171] using fluorescent substrates in reaction mixture of 0.1 M Tris pH 7.25. (B) Effects of metal ions on activity are displayed in bar diagrams using chloride salts of metals: Al^{+3}, Ca^{+2}, Cd^{+2}, Cu^{+2}, Fe^{+2}, K^{+}, Li^{+}, Mg^{+2}, Mn^{+2}, Na^{+} and Zn^{+2} which are supplied at concentration of 1 mM to expression lysate (white bars) and pure lipases (patterned bars). Error bars are showing the deviations of the duplicate measurements.

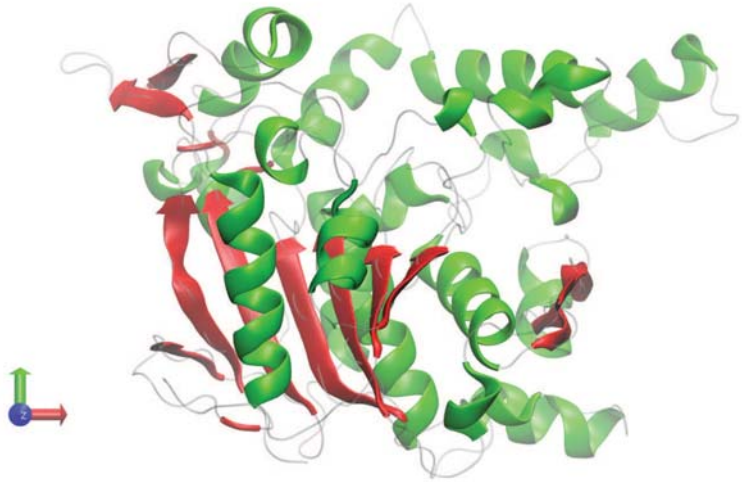

Figure 3.5:
Crystal structure of BTL2. The crystal structure (PDB ID:2W22) is visualized using VMD. Central beta-sheets are colored in red and alpha-helices are shown in green cartoons. The loop regions are transparent grey ribbons.

seven strands of hydrophobic β-sheets packed in central and surrounded by α-helices. Figure 3.5 illustrates 3D structure of BTL2 and the secondary structural elements constituting it.

Active Site and Catalytic Cleft

Positions of two detergent molecules (Triton-X-100) in crystal structure aided in identification of active site cleft. The cleft has dimensions of 14 Å in-depth, 18x25 $Å^2$ in-area and is ovoid in shape [33]. The walls of the cleft are lined with hydrophobic residues (A241, I320, I363, L171, L184, L189, L209, L245, L360, L57, M174, F17, F182, F291, P165, Y30, V172, V175, V188, V234, V295, V321 and V365). These residues play role in stabilization of the acyl chains of TAGs. Moreover the aromatic side chain of the residue F17 cuts the cleft in two. Four binding pockets have been identified in the cleft; an oxyanion hole and three pockets for the different branches of the TAG substrate. The positions of the oxyanion hole and of the three pockets that accommodate the *sn*-1, *sn*-2 and *sn*-3 fatty acid chains is also identified in 2W22. S114 is the catalytic residue whereas the other

two residues (H359, D318) that have role in stabilization of the catalytic attacks made by serine. Serine is found in the motif of A-X-S-X-G, abundant in thermoalkalophilic lipases.

Lid Domain

As with most lipases, another structural element, highly mobile, is also located at the entrance of the active cleft. This element is referred as lid. Two short α-helices (α6 and α7), composed of residues from comprising the lid structure, are found to be linked to core with high flexibility. The lid is opened due to specific interactions of the substrate with the lid helices, a process called interfacial activation. Compared to the closed-form, the C-termini of both helices move about 20 Å away from the entrance of the active cleft [156]. Two structures corresponding to BTL2 is used in Figure 3.6 to see the differences in the open-lid (PDB ID: 2W22) with respect to its closed form (PDB ID: 1KU0). In the closed form, two helices compose the lid structure. The second helix (Figure 3.6) moves laterally away from the cleft opening. This movement is simply rolling of the second α-helix to move its center of mass about 20 Å. Although this helix does not show any change in its structure and composition, its movement causes disruption of the first helix in lid. The first helix in the closed conformation is broken into two smaller α-helices which make the catalytic serine available for substrates (Figure 3.6). Moreover, the last α-helix in the lid changes its conformation to a 3_{10} helix which further contributes to the enlargement of the cleft opening.

Zinc Domain

Extra domain for Zn-coordination, common among lipase family 1.5, is also found in the structure [156]. The open conformation and previously identified closed-conformation of *Bacillus stearothermophilus* lipase (Figure 3.2 do not show significant distinctions in their zinc domains [28]. The extra domain is assumed to account for the increased thermosta-

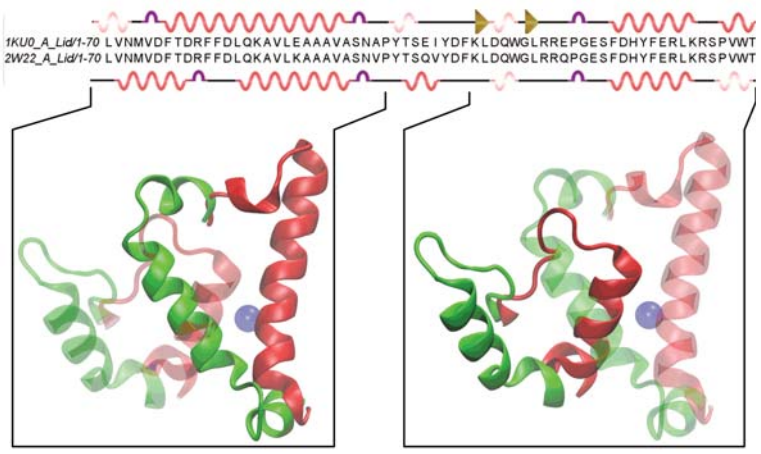

Figure 3.6:
Lid Conformations. Multiple alignment of lid region and secondary structure of lid in the open and closed conformation. Top sequence and secondary structure of the lid is for closed conformation (1KU0) and bottom is for the open state (2W22). The changes in the secondary structure of the lid domain is also given in top view representation of lid helices (green: open-conformation, red: closed, blue sphere: Cα of catalytic serine). The same legend in Figure 3.2 applies here for the secondary structure representations given on top and bottom of the sequence.

bility in this family of lipases [156].

3.3.5 Potential Applications of BTL2

With respect to the chain length selectivity, BTL2 stands out from others as it has the highest specific activity for the short-chain fatty acid C4 TAG compared to other chain-lengths [155, 162]. BTL2 is a thermostable lipase belonging to the Lipase Family 1.5 and shows characteristics of this family [63, 90, 160, 164]. Thermostable lipases like BTL2 are of high importance in industry because of their tolerance to extreme conditions faced in industrial processes [10, 18, 173–175]. In this sense, BTL2 could be an effective alternative to its biocatalyst counterparts in industry [176–182]. Moreover, members of the Lipase Family 1.5 have 95% sequence identity, and hence any finding on BTL2 is likely to be valid for the whole family [160].

Considering previously mentioned lipase applications in various industrial sectors (Ta-

ble 1.1 and 1.2), the native BTL2 is appropriate both for enrichment of long-chain triglycerides and production of short chain fatty acids as it has highest specific activity toward tributyrin and may be used in dairy industry in preparation of cheese flavors as well as milk fat production. Moreover, BTL2 is shown to act as a biocatalyst for these hydrolytic reactions with different chiral substrates [163, 183], which makes BTL2 and its catalytic properties important for industry. The native properties found in BTL2 may be improved through protein engineering to be applied in such sectors. Moreover, since BTL2 is thermostable lipase, it would make the industrial reaction more efficient when compared to its mesophilic counterparts [93].

4 RATIONAL DESIGN OF THE CHAIN LENGTH SELECTIVITY

4.1 Summary

Rational design for *Bacillus thermocatenulatus*lipase (BTL2) is performed to lower the activation barrier for hydrolysis of short chain substrates. In this design, two computational models for the enzyme substrate complex is used for tributyrin (C4) and tricaprylin (C8) which are generated through docking and molecular dynamics simulations. These ES complexes are employed in steered molecular dynamics (SMD) with Jarzynski's equality to estimate their relative binding free energies. Potential mutation sites for modifying the chain-length selectivity of BTL2 are found by inspecting the SMD trajectories and fine-tuning the volume and hydrophobicity of the cleft. The implementation of the SMD method allows the appraisal of the binding free energy of the ES complex at the interface, which might be an experimental challenge. In the experimental section, the steady state specific activities towards C4 and C8 TAGs are determined and combined with the SMD findings to profile Gibbs energy diagrams for C4 and C8 hydrolysis in BTL2 mutants. Here, the implementation of the computational method not only provides us with the binding affinity of ES complex, but also enables us to inspect the dynamics of the ES complex which is useful for rational design purposes. The combined approach described in this study can be used for other lipases and enzymes to investigate their complex reaction mechanisms via single molecule models and also, to design enzymes with altered substrate-selectivity.

4.2 Methods

4.2.1 Structure Preparation

The crystal structure captured in the open conformation (PDBID: 2W22) is used for modeling *Bacillus thermocatenulatus* lipase and its TAG substrates. TAG molecules with a glycerol as the head group tailed with three chains at lengths of tributyrin (C4), tricaprylin (C8) or trilaurate (C12) are used as the ligands. Autopsf plugin in VMD is used to generate structure files using the CHARMM topology parameters of three ligands (given in Appendix A.3. Mutator plugin is used for *in-silico* mutagenesis of lipase structures.

4.2.2 Docking

TAG molecules at the chain-lengths of C4, C8 and C12 are recruited to automated docking suite [115] as ligands. 2W22, the open-conformation, is used as the receptor molecule by eliminating two Triton-X 100 molecules and other ligands in the structure. The ligand is left flexible whereas receptor is rigid during docking. C4 and C8 ligands possess a torsion tree of 14 and 26 rotatable bonds, respectively (Figure 4.1). A total of 10 local searches via Lamarckian Genetic Algorithm (GA) are carried out for the global minimum of the bound complex. Each search consists of a population of 200 members. The search space is contained the catalytic cleft and it is further inspected to include whole opening of this cleft. The exhaustiveness level is set to 128 for Vina calculations in the energy range of 10 kcal/mol. The complex configurations are obtained after the analyses of the docking outcomes by means of poses and scores of poses. The final configurations of 2W22-C4 and 2W22-C8 complexes are decided according to two critical parameters such as the catalytic distance and orientation. Unless the complex configurations secure side-chain orientations and distances of the side chain hydroxyl of S114 and the ester oxygen of the bond joining glycerol to fatty acid of TAG substrate is replicated as in the crystal [33], they are ignored.

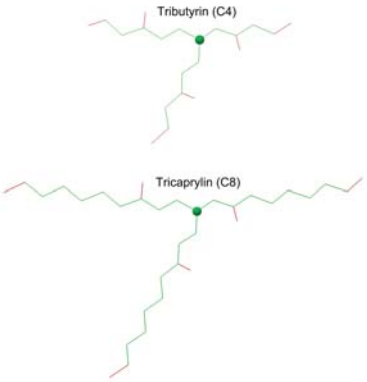

Figure 4.1:
Torsion Tree of C4 and C8 Ligands (Green: rotatable, red: non-rotatable).

The complex molecules having catalytic distance lower than 5 are subsequently simulated in MD. Due to relatively larger size of C12 ligand, docking procedures used for C4 and C8 ligands are not particularly successful for C12 ligand. To replicate crystal parameters in 2W22-C12 complex as well, the coordinates of 2W22-C8 complex are used to generate structure files by exchanging the CHARMM topology parameters of C8 to C12.

4.2.3 Molecular Dynamics

Thre systems containing 2W22 structure as the receptor and C4, C8 or C12 as the ligand are generatured to compose of 55,000, 57,000 and 58,000 atoms, resepctively. The complex systems are placed water box with dimensions of 72x83x92, 82x83x97 and 82x85x95 $Å^3$, respectively. NAMD with the CHARMM22 parameters is used for molecular simulations which are carried out in water box under natural ionic conditions which is 100 mM NaCl as previously defined for thermoalkalophilic lipases [184]. An NpT ensemble is used in MD simulations with periodic boundary conditions. The long-range Coulomb interactions are computed using the particle-mesh Ewald algorithm. Pressure and temperature is kept constant at 1 atm and 300, respectively using the Langevin pressure and temperature coupling. A time step of 2 fs is used in all MD simulations. Following a 1000 step energy minimization, all systems are simulated for nanosecond time scales.

4.2.4 Steered Molecular Dynamics

The starting structure files for SMD simulations are taken from the previous MD simulations. Each starting configuration is equilibrated in the absence of external forces before the SMD method is used to pull the ligand from the receptors cleft. The cleft axis is chosen as the reaction coordinate which is aligned with the z axis for convenience. Despite the fact that the cleft is buried 14 Å deep into the protein core, a straight path can still be drawn from the catalytic triad to the cleft opening, which allows use of a unidirectional force in SMD simulations. The third carbons of all acyl chains in $C4$ and the seventh carbons of all acyl chains in $C8$ are selected as the SMD atoms where the pulling forces are applied. A harmonic force with a constant $k = 20$ kcal/mol.Å^{-2} is cantilevered to a dummy atom placed at the center of mass of the three SMD atoms of each ligand, which is distributed among the three carbon atoms. The ligand is pulled at a constant speed v until it is completely dissociated from the enzyme. Pulling speeds from 20 to 2.5 Å/ns are considered, and the speed of 5 Å/ns is found to yield optimum results for accuracy vs. computing time. The total force experienced by the ligand is given by,

$$F(t) = k[vt - z(t)] \tag{4.1}$$

where z is the displacement of the dummy atom from the initial time to time t. The pulling lasted for 6, 7 and 8 for the ligands $C4$, $C8$ and $C12$, respectively. Ten independent SMD simulations are performed for each ligand. The work function for each SMD trajectory is calculated from the line integral of the force in Equation , and the potential of mean force (PMF) is determined from the individual work functions using Jarzynskis equality [109],

$$e^{\frac{-G_S}{k_B T}} = \langle e^{\frac{-W}{k_B T}} \rangle \tag{4.2}$$

where ΔG_S is free energy difference, W is the work function, k_B is the Boltzmann constant, and T is temperature. The same procedure is used in calculation of the PMFs of ligands bound to the mutant enzymes.

4.2.5 *in-silico* Mutagenesis

MD and SMD trajectories are used to predict mutations in the catalytic cleft of 2W22. Eight different mutants are generated using 2W22-C4 and eight for 2W22-C8 complexes. C12 ligand is not employed in mutant structures. Mutations are chosen so as to increase or decrease the cleft volume available to the acyl chains of the C8 TAG. In order not to disturb the binding of C4, only the residues that interact with the 5th and further carbon atoms of the C8 acyl chains are considered for mutation. For this purpose, pair distances of the C8 ligands 5^{th} carbons with the side chain atoms of the selected lipase residues are measured in initial SMD runs. The residues that have yielded pair distances less than 10 with the 5^{th} carbon of the C8 ligand are considered in site-directed mutagenesis. Four different mutants with single amino acid changes (L57F, V175F, I320F and L360F) are considered. Two residues from the *sn*-1 pocket (V175 and I320), and one residue from the *sn*-3 pocket (F17) are also mutated to alanine to further investigate the effect of volume change in the binding. The side chains of all of the selected residues point towards the cavity and do not have any interactions with the protein backbone. The active site residues residing in the oxyanion hole are excluded in mutagenesis studies to avoid substantial decrease in lipase activity.

All of the mutants systems are energy minimized in 1000 steps and employed in MD simulations. To investigate chain-length selectivity, 8 mutants of 2W22-C4 and 2W22-C8 complexes are simulated for 1 ns in the absence of external forces and the starting configurations for SMD method are selected from the last 0.1 ns. For the mutant systems, L57F, V175F, I320F, and L360F, the configurations at the end of 1 ns have resulted in a

wrong conformation of the C8 ligand (e.g., the catalytic pairings observed in the crystal structure is lost and the position of C8 is shifted along the z-axis). Therefore restraints are introduced on the C8 ligand for these mutant systems. The head group carbons of the ligands are initially restrained with a harmonic force of 10 kcal/mol.$Å^{-2}$, which is reduced to 0 in five equal steps over the course of a 500 ps MD simulations. With this procedure, complex structures consistent with the crystal structure are obtained. Again each mutant system is simulated without restraints for 1 ns to provide the initial configurations for the SMD simulations. 5 pulling events with constant speed of 5 Å/ns are carried out for each mutant in SMD and PMF is profiled for each mutant, similarly with Jarzynski's Equality.

4.2.6 Molecular Cloning

A 1,167-bp DNA fragment corresponding to the *Bacillus thermocatenulatus* lipase gene (BTL2) is amplified from the mature lipase clone (pPICZα A-BTL2) DNA using the primer set containing ligation independent cloning (LIC) sites; (**bold**) for forward (F_ BTL2_ LIC: 5- **TACTTCCAATCCAATGAA** GCGGCATCCCCACGCG -3 and for reverse (R_ BTL2_ LIC: 5- **TTATCCACTTCCAATGAA** AGGCCGCAAACTCGCCAA -3). PCR conditions are given in Table 4.1. 4 μg of the expression vector, pMCSG7 is linearized using *Ssp*I restriction enzyme (see Appendix A.4for vector maps). Linear vector and PCR product are electrophoreted at 100 V in 1.2% agarose gels using tris borate edta (TBE) buffer system for 30 minutes. Both fragments are extracted from agarose gel and treated with T4 DNA Polymerase. The exonuclease activity of T4 DNA polymerase is terminated using excess concentration of dGTP for vector and dCTP for PCR product depending on the given LIC sequences. The reaction is lasted at 20^oC for 50 minutes and heat-inactivated at 70^oC for 20 minutes. T4 DNA polymerase treated DNA are subsequently extracted to aqueous phase using phenol/chloroform and precipitated in cold-ethanol. Vector and PCR product are solubilized in 10 μl of distilled water and annealed

Table 4.1: PCR Cycle Profile for BTL2 Amplification.

Step	^oC	min	Cycle
Initial Denaturation	94	3	1
Denaturation	94	0.5	
Annealing	52	0.5	35
Extension	72	1	
Final Extension	72	7	1
Hold	4		

at 23^{o}C for overnight. Chemically competent *E. coli* XL1-Blue cells are prepared according to ref. [185]. All of the reaction mixture is transformed to *E. coli* XL-blue cells via chemical transformation [186] Colony PCR reactions are set using BTL2_LIC primer pairs and given with cycle profile given in Table 4.1 for two selected colonies. All of the amplifications are size separated in 1.2% agarose gels using GeneRuler^{TM} 1 kb DNA Ladder SM0311 (Fermentas, see Appendix A.1 for the marker legends). Plasmid purifications are performed from two colony-PCR positive clones and purified plasmids are digested with *Eco*RI and *Nde*I to confirm BTL2 gene. The plasmid purifications are done using Qiagen Plasmid Purification kit following the manual. The restriction digestion with *Eco*RI and *Nde*I is performed as recommended by the supplier (Fermentas). Plasmids are sequenced using T7 Terminator (5- GCTAGTTATTGCTCAGCGG -3) and T7 Promoter (5-TAATACGACTCACTATAGGG-3) by Molecular Cloning Laboratories (MCLAB).

4.2.7 Site-directed Mutagenesis

7 single mutations and one double mutation are generated using Overlap Extension PCR (OE-PCR). The mutations and corresponding primers are listed in Table 4.2. For alanine substiution, GCG triplet, for phenylalanine, TTT is preferred obeying the codon preference displayed by *E. coli* [187]. For each mutant, two amplification reactions are set: (i) using primers: F_BTL2_LIC as forward and mutant primers as reverse (ii) mutant primers as forward and R_BTL2_LIC as reverse -a schematic representation for this reaction is given

in Figure 2.2. These fragments are extracted from 1.2% agarose gel and used to amplify full-length fragment in OE-PCR. The first 15 cycles in this PCR are performed without any primers. After 15 cycles, BTL2_LIC primer pairs are added into reaction mixture. The amplification conditions for sequential PCRs explained here is given in Table 4.3.

Table 4.2: Mutagenesis Primers.

Mutation	Direction, 5-3	Sequence, Mutant Codon, in bold
F17A	forward	GTGCTTCTCCATGGG**GCG**ACAGGATGGGGGCG
	reverse	CGCCCCCATCCTGT**CGC**CCCATGGAGAAGCAC
L57F	forward	CGCTGGCGGTCGGACCGT**TTT**CGAGCAACTGGGACCG
	reverse	CGGTCCCAGTTGCTCG**AAA**ACGGTCCGACCGCCAGCG
V175A	forward	GACGACACTTGTCAATATG**GCG**GATTTCACTGATCGCTT
	reverse	AAGCGATCAGTGAAATC**CGC**CATATTGACAAGTGTCGTC
V175F	forward	GACGACACTTGTCAATATG**TTT**GATTTCACTGATCGCTT
	reverse	AAGCGATCAGTGAAATC**AAA**CATATTGACAAGTGTCGTC
F177A	forward	CACTTGTCAATATGGTCGAT**GCG**ACTGATCGCTTCTTTGACC
	reverse	GGTCAAAGAAGCGATCAGT**CGC**ATCGACCATATTGACAAGTGT
I320A	forward	GCTGGAGAACGATGGC**GCG**GTCAACACGGTTTCG
	reverse	CGAAACCGTGTTGACCG**CGC**CATCGTTCTCCAGC
I320F	forward	GCTGGAGAACGATGGC**TTT**GTCAACACGGTTTC
	reverse	GAAACCGTGTTGAC**AAA**GCCATCGTTCTCCAGC
L360F	forward	TACAACGTCGACCAT**TTT**GAAGTGATCGGCGTTG
	reverse	CAACGCCGATCACTTCA**AAA**TGGTCGACGTTGTA

Table 4.3: PCR Cycle Profile for Site-directed Mutagenesis and OE-PCR.

	1^{st} **Round**		2^{nd} **Round**		**Cycle**
Step	^oC	min	^oC	min	
Initial Denaturation	94	3	94	3	1
Denaturation	94	0.5	94	0.5	
Annealing	65	1	53	0.5	35
Extension	72	1	72	1	
Final Extension	72	7	72	7	1
Hold	4		4		

The mutant fragments are cloned into pMCSG-7 backbone using same cloning procedure given described for BTL2. The double mutant is amplified after construction of I320F mutant. The PCR reaction for the double mutant is used I320F clone -sequence confirmed- as the template and L360F mutant primers given in Table 4.2.

4.2.8 Lipase Expressions

Sequence confirmed clones are transformed to Shuffle *E. coli* cells via chemical transformation and single colonies are selected for expression studies. Each clone (wild-type, F17A, V57F, V175F, V175A, I320F, I320A, L360F and I320F/L360F) are plated on LB-rhodamine plates which are prepared with slight modifications to the ref. [188] to check the lipase activity. All mutants are also expressed in liquid culture using 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) in LB. The expressions are lasted out for six hours and sampled once at the fourth hour. The cells are harvested by centrifugation (maximum speed on table top microcentrifuge for 5 min) and lysed using B-PER (Thermoscientific). B-PER solubilized samples are fractioned by centrifugation (maximum speed on table top microcentrifuge for 10 min). Lipase activity of soluble fractions is determined using fluorescent substrates, 4MU-C4 and 4MU-C8 in 0.1 M Tris-Cl at pH 7.25. Soluble fractions are analyzed by SDS-PAGE (sodium dodecyl polyacrylamide electrophoresis) and visualized after coomassie staining. Time-step for the highest expression is determined for each

clone and another round of expression of is carried out under same conditions with larger volumes. The expressions are terminated by centrifugation at 10,000 rpm for 15 min at 4^{o}C . Unless the samples are purified subsequently, the pellets are frozen at -20^{o}C immediately and kept at -80^{o}C.

4.2.9 Lipase Purifications

Nickel coated beads of poly-histidine tagged proteins (GenScript) are used for affinity purification of wild-type and mutants. Sodium phosphate buffer at 20 mM provided with 50 mM Imidazole during binding and 500 mM Imidazole during elution steps. 10 kDa Filter Concentrators (Millipore) are used for buffer-exchange of 20 mM sodium phosphate pH 8.0 against 500 mM Imidazole. The quantities of the purified lipases are measured in Bradford protein assays [189] and sodium dodecyl polyacrylamide electrophoresis (SDS-PAGE) is used to test the purity of the lipases. Stacking is made using 4% polyacrylamide and size separatiion is made under 12.5%.

4.2.10 Lipase Assays

Lipase activity is determined in two different assays. In titrimetric assay, triacylglycerols (TAGs) are used as substrates of lipase reactions. The liberated fatty acids are quantified by direct titration with sodium hydroxide to thymolphthalein end point as described [190]. In fluorometric assay, 4-MU based substrates are used.

Titrimetric Assays

Lipase assays at saturating conditions are performed to measure steady state activities. Specific activities of BTL2 and mutants are measured in enzyme assays using emulsifications of C4 and C8 TAGs at final concentrations of 100 and 150 mM with 5% (w/v) gum arabic as stabilizer. Lipase activity at pH 8.0 is determined using purified lipases by direct

titration of fatty acids as described [190] and 1 unit is defined as the amount of enzyme to liberate 1 μmole fatty acid in 1 minute. The reactions are performed at room temperature ($\sim 300K$) to be consistent with computational models and at 55^oC. Because 6-8% change in specific activities are observed for 100 mM substrate concentration, specific activities of lipases are calculated at 150 mM substrate concentration and expressed in Units.mg^{-1} of pure lipase. The assays are repeated for two times to test its reproducibility. Lipase activity is also measured at 55^oC at pH 8.0 using same titration procedure for all lipases. Assuming all lipase bound to the lipid interface of substrate emulsion, the specific activity values reflect a measure from steady state and the fold changes in specific activities of mutants with respect to wild-type correspond to the ratio of $k_{cat_{mt}}/k_{cat_{wt}}$, which affects the binding free energy of ES‡ complex ($\Delta\Delta G^{\ddagger}$) with a logarithmic factor [191] The Gibbs energy diagrams for mutants are profiled using two terms, $\Delta\Delta G_S$ and $\Delta\Delta G_{\ddagger}$. $\Delta\Delta G_S$ is the difference in the binding free energy of mutants relative to wild-type and is calculated by SMD. $\Delta\Delta G^{\ddagger}$ is calculated by [192, 193],

$$\Delta\Delta G_T^{\ddagger} = -k_b T \ln(k_{cat_{mt}}/k_{cat_{wt}}) \tag{4.3}$$

Fluorescent Assays

For fluorescence assays, lipase activity is measured in a 96-well black micro-titer plate using 4MU-butyrate and 4MU-caprylate as substrates. Cell-free expression medium is assayed using different substrate concentrations in reaction medium of 100 mM Tris-Cl at pH 7.25. Gemini XS (Molecular Devices) is used to measure 4MU-fluorescence using an excitation wavelength of 355 nm and an emission wavelength of 460 nm in a kinetic fashion. The measurements are made at every minute for 1 hour. All measurements are performed in duplicates and initial velocities are calculated using SoftMaxPro Software (Molecular Devices). Relative Fluorescence Unit obtained from fluorometer is converted to

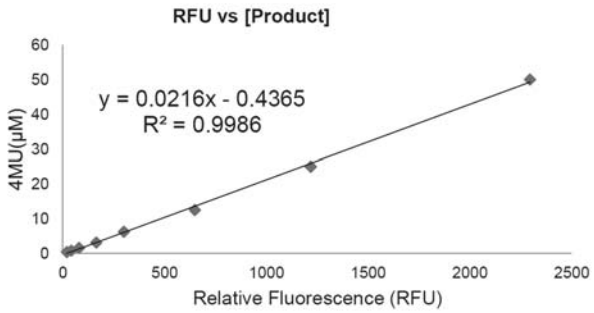

Figure 4.2: Standard Curve for RFU and 4-MU

4MU units according to the linear relationship shown in Figurefig4.2. Kinetic parameters are obtained from Lineweaver-Burk plots.

4.2.11 Thermostability Analysis

Purified lipases are used in fluorescent assays to profile thermostability by quantifying the residual activity of lipases after incubation for 30 minutes at temperatures of 60^{o}C, 65^{o}C, 70^{o}C, 75^{o}C and 80^{o}C. The reaction is performed in a mixture of 250 μM substrate, 0.1 mM Tris-Cl pH 7.25 at room temperature. The percent activity is calculated by setting 60^{o}C to 100%.

4.3 Results and Discussion

4.3.1 Characterization of Complex Molecules

Complex configurations are delineated to validate bound ligand in BTL2 cleft. Initially using MD results preceding SMD, the cleft region is identified to comprise of four binding pockets, three pockets for the acyl chains of TAG (*sn*-1, *sn*-2, sn-3) and an oxyanion hole for head group. The catalytic cleft of BTL2 is formed to unite *sn*-1 and *sn*-3 pockets that resembles boomerang i.e. two large crevices for three acyl chains in Figure 4.3.

The catalytic triad of BTL2 is formed by S114, D318 and H359. The investigations of the computational models employed in this section ensure correct orientations of the

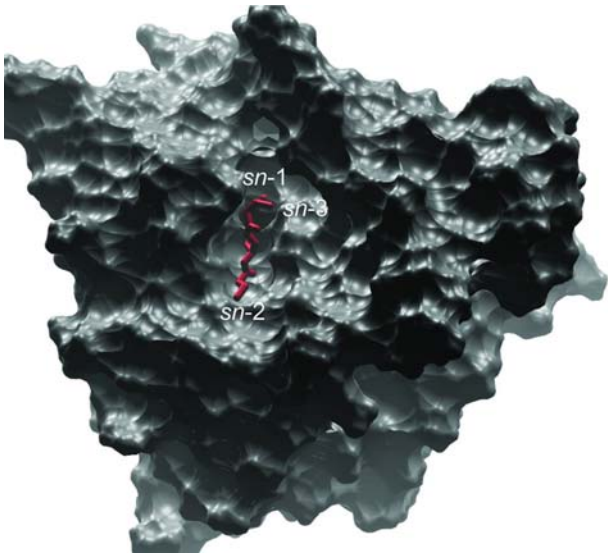

Figure 4.3:
Top-view of the BTL2-C8 complex. BTL2 surface is shown in grey and C8 ligand is in red stick model. To depict the boomerang shape of the cleft, *sn*-1, *sn*-2 and *sn*-3 acyl chains are labeled.

catalytic triad residues (S114, D318 and H359) and ligands (Figure4.4). Decreased fluctuations of active site residues (less than 1 Å) also suggest tight packing of catalytic triad as shown in Figure 4.5 A. Moreover, the head group atoms of *sn*-1 chain is shown to be less mobile compared to acyl chain atoms in Figure 4.5 B. Decreased mobility in the active site as well as the scissile fatty acid of both ligands reflects bound configuration for ES complex for TAG hydrolysis.

4.3.2 PMF from SMD simulations

The pulling speed was another critical parameter that the accuracy of the free energy calculations from SMD trajectories is valid only at very slow speeds [151]. However required computer-power for slower rates exponentially increases. Therefore, a compromise is essential between accuracy and cost. A study like ours would least likely to suffer from such compromise. In our case, ranking of binding free energies of three different TAG molecules to wild-type and mutant receptors was aimed. The speeds; 10 Å/ns, 5 Å/ns and

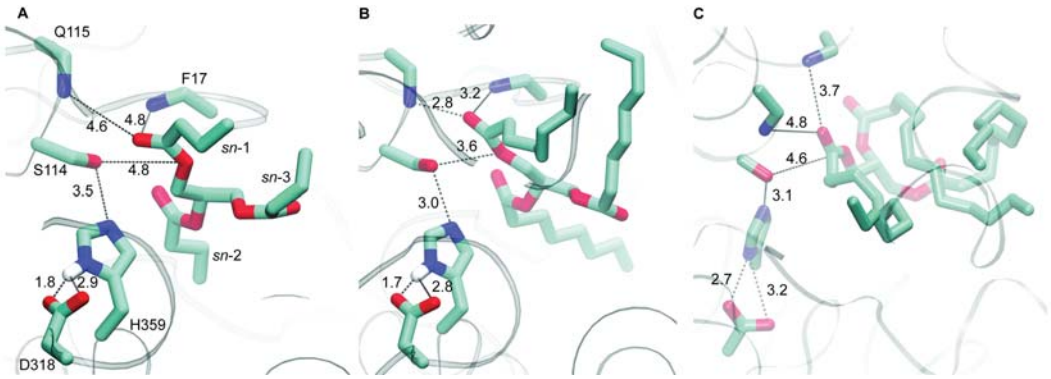

Figure 4.4:
ES Complex for C4, C8 and C12. (A) C4, (B) C8 and (C) C12 TAGs are shown in stick models (C: cyan, O: red, N: blue) along with active site residues to mark the catalytic triad (S114, D318 and H359) couplings. Distances between *sn*-1 carboxyl oxygen and S114 side-chain hydrogen and between catalytic triad residues are shown with dashed lines are also given in . The protonation of histidine in the active site of the enzyme-substrate complex is replicated as shown here. For (C), histidine hydrogen are not shown.

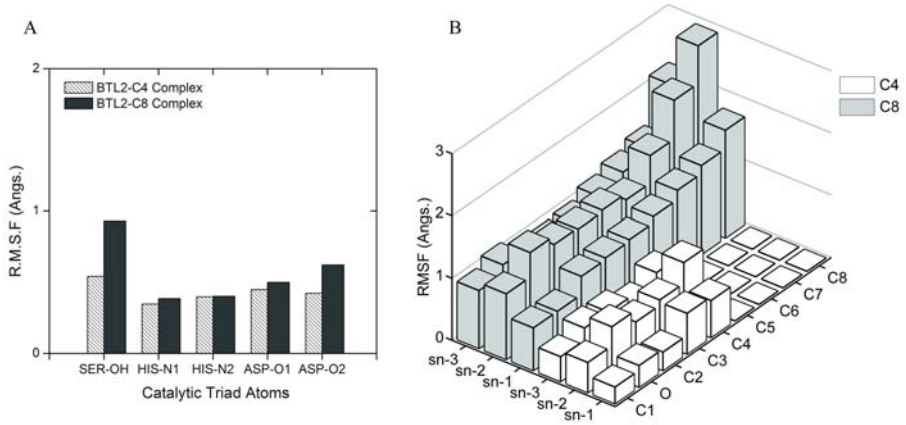

Figure 4.5:
Fluctuations of Ligands and Active Site Residues. RMS fluctuations of the active site residues (A) are plotted for 1 ns of equilibration with light bars for the BTL2-C4 complex, and dark bars for BTL2-C8 complex. The fluctuations for acyl chain atoms of ligands (B) are plotted using white bars for C4 and grey for C8.

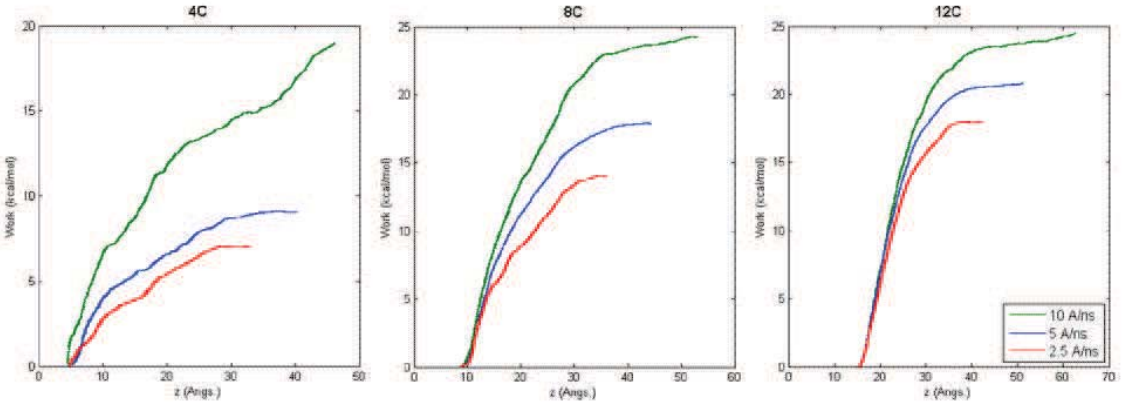

Figure 4.6:
Pulling Speeds for SMD. Work functions plotted against the reaction coordinate for C4, C8 and C12 ligand at pulling speeds of 10 (green), 5 (blue) and 2.5 Å/ns (red).

2.5 Å/ns were used to calculate work values of attempted pullings for three ligands (Figure 4.6). To save computer power, 5 Å/ns was found optimum for SMD.

The work functions obtained from ten SMD trajectories for the C4, C8 and C12 ligands are shown in Figure 4.7. The PMF profiles calculated from these work functions using Jarzynskis equality are indicated with black lines in the same figure, from which the relative binding free energies for the C4, C8 and C12 TAGs are found as -9.8, -17.9 and 21.2 kcal/mol, respectively. These results indicate that using ten samples is sufficient for the purpose of ranking of all ligands.

It is noted that the catalytic cleft of BTL2 is 14 Å in-depth [33], but a flat PMF shape in Figure 4.7 becomes evident only after a pulling distance of 25 Å. To understand why the dissociation of the ligands from BTL2 is delayed for another 10 Å, the individual SMD trajectories are inspected and it is found that despite the application of a unidirectional force in the z-direction, in some of the pulling events, the ligand bends towards the lid after coming out of the cleft and starts interacting with the hydrophobic residues on the lid surface. This happens 3 out of 10 cases for C4, 6 out of 10 cases for C8, 5 out of 10 cases for C12. One example of such a bent path for the C4 ligand is shown in Figure 4.8. In reality, the lid may interact with several TAGs simultaneously, and it is difficult to estimate

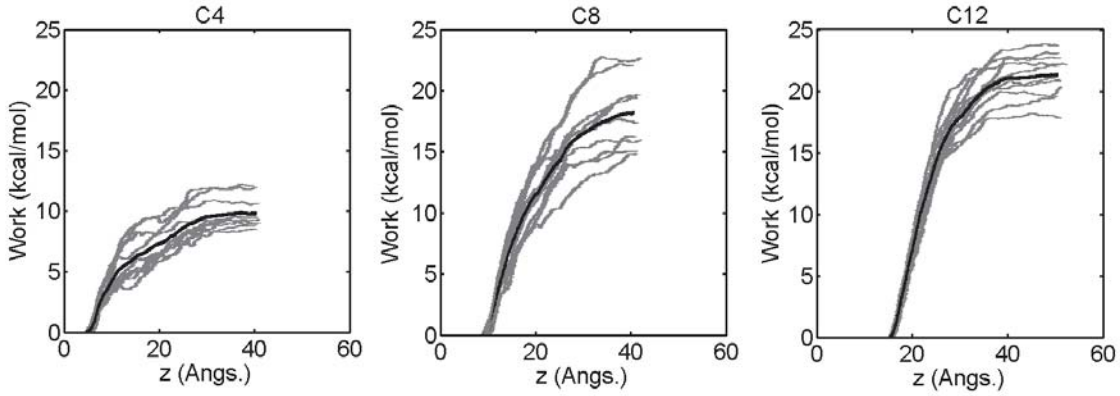

Figure 4.7:
PMF for C4, C8 and C12 TAG binding. Work functions plotted against the reaction coordinate for C4, C8 and C12 ligand. The reaction coordinate is chosen as the cleft axis and the initial reaction coordinates of both ligands are aligned in the catalytic cleft according to their center of masses. PMF profiles of C4, C8 and C12 binding are constructed from 10 work functions and shown in black.

the actual free energy of lid binding from SMD simulations. In any case, the binding of single TAG to the catalytic cleft which corresponds to the affinity of ES complex at the interface is investigated here. Thus, the cleft binding of the ligands from lid binding is distinguished from each other. The free energy of cleft binding of TAG molecules from the first 15 Å of the SMD extension is determined as -7.2, -15.2 and -18.0 kcal/mol for the C4, C8 and C12 ligands, respectively. Comparing with the binding free energy obtained from the full extension, the free energy for the lid binding can be estimated as about -2.6 kcal/mol for all TAGs. Thus, considering only the free energies of cleft binding does not change ranking of the ligands.

All ligands have the same configurations for binding of the head group atoms to the catalytic triad (Figure 4.4) but the C8 ligand has four extra hydrocarbons in the binding pocket involved in hydrophobic interactions. This explains the larger affinity of C8 for BTL2 compared to C4, similarly C12 has increased affinity for the cleft compared to smaller TAGs. Because the binding affinity is overestimated in SMD calculations, especially for larger ligands [148, 149], the binding free energy difference between C8 and C4 is taken as a

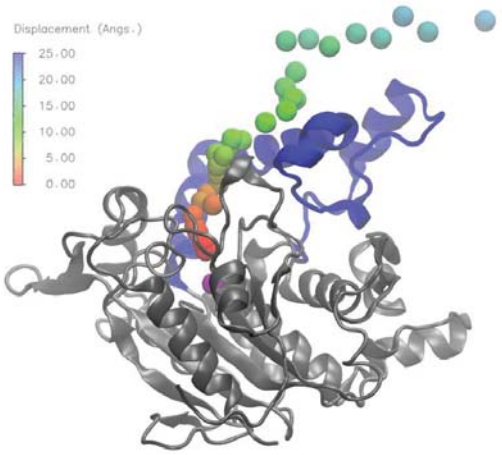

Figure 4.8:
The Bent-exit of C4 from Catalytic Cleft. An example dissociation of C4 ligand from catalytic cleft (i.e. pulling upwards) is illustrated. Displacement of COM of C4 ligand is shown by color spheres and RGB color scheme for displacement of C4's COM is scaled in bar is used to color the COM spheres with respect to the corresponding distance from the catalytic cleft during pulling. Lid helices is shown in blue to clarify the bending towards them of C4 ligand's COM towards these helices. CA atom of catalytic serine is represented as magenta sphere to assess the cleft depth.

qualitative estimate rather than a quantitative prediction.

4.3.3 Catalytic Cleft of BTL2 and Selection of Mutants

Among three ligands, C4 and C8 TAGs are selected to investigate the chain-length selectivity of BTL2 and C12 is eliminated in mutant analysis. The reason for not considering larger molecules than C8 TAG (e.g., C12) is that the accuracy of the SMD method in free energy calculations is reasonably good for small molecules but it becomes increasingly unreliable for larger ligands [128, 148, 194] Moreover, BTL2 shows the highest activity for C4 TAG followed by C8 which has half the activity of C4 [162] Thus, C4 and C8 TAGs are considered as an optimum ligand pair to assess the chain-length selectivity of BTL2.

Chain-length selectivity is determined by the volume and the hydrophobicity of catalytic cleft. By inspecting the SMD trajectories of TAGs, one can find the critical BTL2 residues involved in binding of the acyl chains. Positions of the critical residues identified

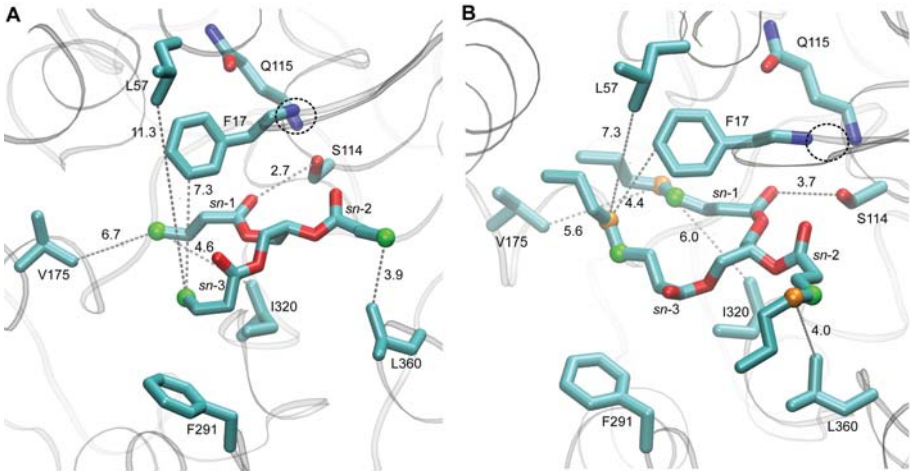

Figure 4.9:
Mutant Selections. The initial structures of SMD method is shown for both ligands, (A) C4 and (B) C8. The pulling direction is towards lower left corner of the Figs. The side chains of the ligands and the selected residues are shown by stick models (C: cyan, O: red and N: blue). The respective distances between the selected side chain atoms of the residues and ligand atoms are given in units of Angstrom. These distances are measured for the fourth and fifth carbons of ligand are represented by green and orange spheres, respectively. Catalytic distance is also shown for S114 residue along with the orientation of the double bonded oxygen of *sn*-1 chain towards main chain amide groups of F17 and Q115 residues that form the oxyanion hole (in dashed circle).

from such a search are shown in Figure 4.9 for the C4 and C8 ligands. Because BTL2 is regio-specific to *sn*-1 and *sn*-3 chains as most of the extracellular lipases [71], the catalytic cleft is formed to unite the binding pockets for *sn*-1 and *sn*-3 in one large cavity. This large cavity is lined with hydrophobic residues that interact with the *sn*-1 chain on one side and *sn*-3 on the other side. Inspection of this cavity reveals that two phenylalanine residues, F291 and F17, are involved in binding of the *sn*-1 and *sn*-3 chains, respectively, each packing up to four carbon atoms (Figure 4.9). The cavity is surrounded by I320 at the bottom and V175 at the end, which interact with the *sn*-1 chain, and L57 at the top which interacts with the *sn*-3 chain. Although L360 is in the *sn*-2 pocket and not considered to be critical for the catalytic machinery, it does have a close contact with the *sn*-2 acyl chain (Figure 4.9), and therefore it is included in the list of mutations. Conversely, F291, which is in close-proximity of the *sn*-3 chain of both ligands, is not considered for mutagenesis in

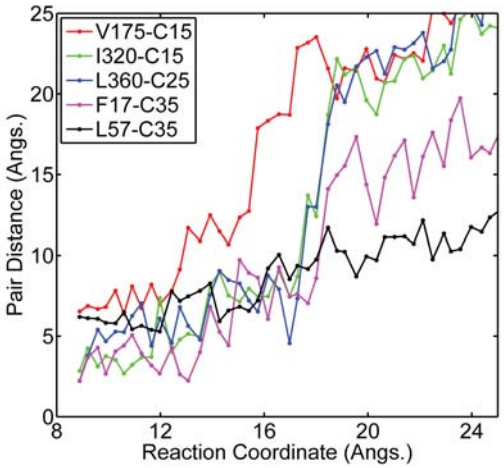

Figure 4.10:
Pair Distances. The measurements of the distances between C8 ligand and the selected residues are plotted against the extension in the reaction coordinate. Here C15, C25, and C35 denote the fifth C atom of the chains *sn*-1, *sn*-2, and *sn*-3, respectively. Average measurements for 5 pulling events for C8 ligand is used to measure the distances. The pairs of atoms used in these distances are illustrated in the given snapshots of initial SMD structures in Figure 4.4

order not to disturb the structural integrity of the catalytic cleft. To confirm the importance of these five residues (F17, L57, V175, I320, and L360) in the binding of the TAGs, we show in Figure 4.10 the pair distances between these residues and the fifth carbons of the C8 ligand as a function of the pulling distance. Persistence of these contacts for several indicates their importance in binding. To increase the selectivity of BTL2 for C4 over C8, we consider mutation of the four aliphatic residues (L57, V175, I320, and L360) to phenylalanine which will reduce the volume available for C8 but not C4, and hence decrease the affinity of C8 relative to C4. An opposite mutation is performed for F17 (i.e., F17A) to investigate the effect of volume increase in the cleft. Similarly two residues from the *sn*-1 pocket (V175 and I320) are mutated to alanine to test our mutagenesis strategy.

A total of 16 mutant complexes (combination of 8 lipases and 2 ligands) are investigated to analyze the effect of the volume adjustments in all three binding pockets of BTL2. Although, the geometry of catalytic triad is highly conserved, that of the binding pockets are highly variable among lipases. The variable composition of the binding pockets gener-

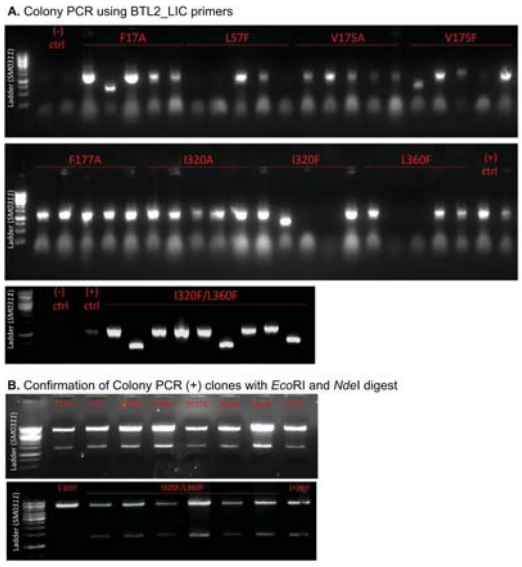

Figure 4.11:
Cloning Confirmations. Negative control ((-) ctrl) is PCR mix only whereas positive control ((+) ctrl) is recombinant pMCSG-7 harboring wild-type BTL2. (A) Colony PCR is performed for 9 clones with BTL2_LIC primer pair and (B) confirmation digests of the pure plasmids are done using *Eco*RI and *Nde*I enzymes. Both analysis are made in agarose gel under 100 V for 30 minutes.

ates an elastic environment for the acyl chains at various chain-lengths [41]. Thus, a large number of mutants is required to clarify the function in such variable environments.

4.3.4 Mutagenesis and Production

Mutagenesis is carried out using OE-PCR strategy which is as effective as kit-based approaches [102] Because usage of ligation independent cloning is proved for efficient handling of large number of clones [195], cloning procedures are ligation-free. Heterologous expressions are performed in *E. coli* and recombinant lipases are tagged with poly histidine at their N-terminus which allows one-step affinity purification.

Eight BTL2 variants -7 single and 1 double mutation- is generated in two PCR steps and subsequently cloned to pMCSG 7 backbone using ligation independent sites. The confirmations of mutant clones are given in Figure 4.11. One of the selected plasmid for

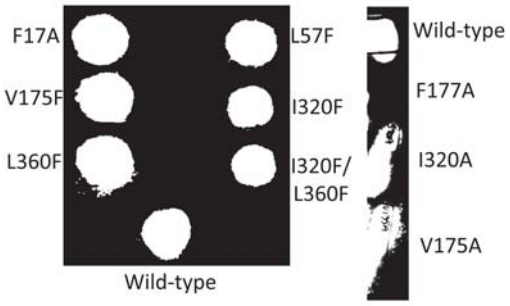

Figure 4.12:
LB-Rhodamine Plate Test for Expressions. After 2 days of incubation on LB-rhodamine, the wild-type and mutant *E. coli* SHuffle clones are visualized under UV light.

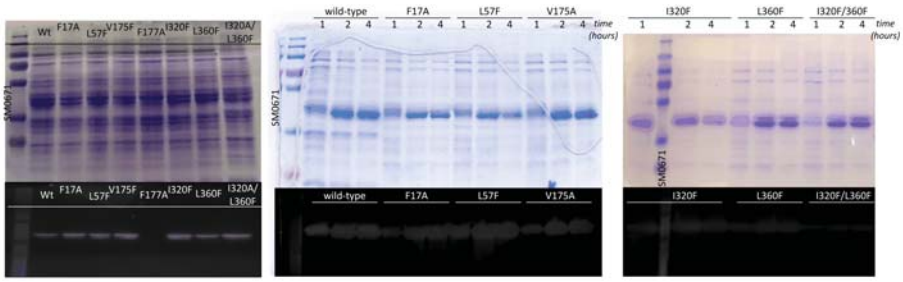

Figure 4.13:
E. coli SHuffle Expressions. Whole cell-fraction from wild-type and mutants, collected after 4 hours, analyzed in 12% (top) SDS-PAGE and (Bottom) zymography using 4MU-butyrate. SDS-PAGE is stained with coomassie and zymogram is performed under UV light.

each mutant is sequenced and used as template of expression studies. Sequence confirmed clones in *E. coli* SHuffle cells are tested against lipase activity and results are given in Figure 4.12. This initial test of heterologous expression is positive for each clone except for F177A (Figure 4.13). Thus, this mutant could not be considered in purification studies (Figure 4.14).

4.3.5 Titrimetric Lipase Assays

The specific activities of BTL2 and its mutants are given in Table 4.4 for both substrates. The specific activity towards C4 chains is enhanced for V175F mutant only whereas C4 activity is decreased for the mutants L57F, V175A, I320A and I320F/L360F. The rest of

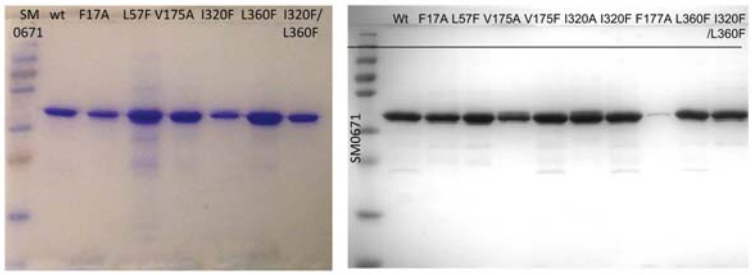

Figure 4.14:
Metal-affinity Purifications. SDS-PAGE result of purification of wild-type and mutant lipases used in this section. First lane in both gels correspond to molecular weight marker (SM0671), The samples are prepared in 0.1 Tris-Cl buffer at pH 8.0.

Table 4.4:

Titrimetric Assay Results. [a]Specific activity (Units.mg^{-1}) of BTL2 and its mutants is determined in titrimetric assays at 25^{o}C using C4 and C8 TAG substrates. The results are averages and standard deviation of two independent assays.

	Specific Activity (U/mg)	
	Tributyrin (C4)	Tricaprylin (C8)
wild-type	583±16	164±19
F17A	518±122	598±26
L57F	369±175	177±93
V175A	399±36	150±39
V175F	1571±257	613±121
I320A	299±31	132±25
I320F	588±133	209±18
L360F	609±103	506±53
I320F/L360F	392±87	157±42

the mutations, F17A, I320F and L360F do not affect specific activity of BTL2 towards C4. On the other hand, F17A, V175F and L360F increase C8 activity while others, L57F, V175A, I320A, I320F and I320F/L360F do not show significant change in the specific activity towards C8 chains.

BTL2 is thermostable lipase that has previously shown its highest activity at high temperatures (50-60^{o}C) and lipase activity is also measured at 55^{o}C (Table 4.5). From these, the double mutant which is the combination of I320F and L360F is found to be inactive at 55^{o}C. The reactions when carried out at ambient temperatures, double mutant is compe-

Table 4.5:

Fold-changes in Activity at Different Temperatures. NA: No activity. The fold-change in the activity od mutants (see Table 4.4) are given with respect to [a] At 25^oC specific activity for BTL2 towards C4: 583 Units.mg^{-1} and towards C8:164 Units.mg^{-1}. [b]At 55^oC the specific activity towards C4: 1153 Units.mg^{-1} and towards C8: 532 Units.mg^{-1}.

Lipases	**25**^oC		55^oC	
	C4	C8	C4	C8
wild-type	1^a	1^a	1^b	1^b
F17A	0.89	3.65	0.43	1.18
L57F	0.63	1.08	0.77	0.75
V175A	0.70	0.81	0.75	0.68
V175F	2.69	3.74	1.02	2.39
I320A	0.50	0.81	0.33	0.30
I320F	1.01	1.27	1.07	1.04
L360F	1.04	3.09	0.97	1.08
I320F/L360F	0.67	0.96	NA	

tent to hydrolyze both of the chains with a slightly increased selectivity towards C8 chains whereas this double mutant is inactivated at higher temperatures like 55^{o}C. This suggests that double mutation in the catalytic cleft is likely to result in the loss of the temperature activity of lipase. Conversion of the aliphatic residues to bulk phenylalanines in the catalytic cleft may have a negative influence on the flexibility of the catalytic cleft [156]. Only, V175F shows differential chain-length selectivity at high temperature. The selectivity shifts to C8 at high temperatures for this mutant. Two explanations previously made by Jeong et. al. can account for this discrepancy observed for V175F which is found in 2 residues away from the N-terminus of the first lid helix: The hydrophobic residues may enhance lid interactions with the insoluble interface at high temperatures and may interfere with the interfacial activation process [156]. In these aspects, V175F can be considered critical for interfacial activation. Another point to support this result can be unchanged V175A selectivity at high temperatures. With respect to the hydrophobic interactions displayed by V175A which is weaker than V175 and much weaker than V175F.

4.3.6 Kinetic Parameters towards 4MU-esters

Fluorescent assays are also performed for all lipases using 4MU-butyrate (C4) or 4MU-caprylate (C8) as substrate to find the chain-length selectivity for acyl esters of fatty acids. The selectivity constant, k_{cat}/K_M is derived from the double reciprocal $1/[V]$ vs. $1/[S]$ (Lineweaver-Burk) plots of the kinetics of 4MU production. 4MU-butyrate (C4) or 4MU-caprylate (C8) are used in fluorescent assays to find the chain-length selectivity towards acyl esters of fatty acids. The most significant changes are observed for I320F for which the selectivity constant is increased by 8-folds for butyric acid methyl ester (C4) (Table 4.6. Experimental quantification of the dissociation constant for ES complex, K_S (when $K_M = K_S$) is trivial for most of the metabolic enzymes having water soluble substrates yet it is a challenge for TAG hydrolysis because the hydrolysis is essentially carried out at water-lipid interfaces. Therefore, kinetic parameters, particularly K_S, for TAG hydrolysis is commonly approximated using single-chain esters of p-nitrophenyl or umbelliferyl that mimics scissile fatty acid (*sn*-1 chain for BTL2) [196]. Although hydrolysis of these substrate analogs is presumptive only for lipase activity, kinetic parameters for triacylglycerol hydrolysis cannot be determined accurately using such analogs. More importantly, Schmidt-Dannert et. al (1996) reported that chain-length selectivity of BTL2 is varied among triacylglycerols and p-nitrophenyl based substrates [197]. Thus these fluorescent assays are performed to see the effects of these mutations on fatty acid methyl esters rather than to validate the computational findings.

Table 4.6: Kinetic Parameters towards Fluorescent Substrates. "Mean ± standard deviation" is given in units of μM for K_M, sec^{-1} for k_{cat}, $sec^{-1}.M^{-1}$ for k_{cat}/K_M. *Statistical significance in the kinetic parameters of mutants with respect to wild-type is marked with $p < 0.01$ after application of students two tailed t- test.

.

Lipases	**Pts**	**Tributyrin (C4)**			**Tricaprylin (C8)**		
		K_M	k_{cat}	k_{cat}/K_M	K_M	k_{cat}	k_{cat}/K_M
BTL2	5	38.4 ±3.5	0.0014 ±0.0002	37.3	102.7 ±16.3	0.016 ±0.005	158.7
F17A	2	37.0 ±6.7	*0.0034 ±0.0005	91.4	*217.8 ±26.9	0.030 ±0.006	137.7
L57F	2	40.8 ±9.1	*0.0032 ±0.0006	78.9	149.3 ±32.1	0.033 ±0.009	200.9
V175F	2	34.2 ±6.2	*0.0032 ±0.0008	94.0	101.2 ±10.2	0.033 ±0.004	296.4
I320F	5	*27.1 ±2.1	*0.0080 ±0.0009	297.8	*373.5 ±69.1	0.064 ±0.006	171.4
L360F	3	31.2 ±6.4	0.0017 ±0.0003	53.4	68.1 ±15.1	0.017 ±0.002	245.2
I320F/L360F	3	44.4 ±2.1	0.0021 ±0.0003	48.8	35.5 ±10.1	0.013 ±0.001	363.4

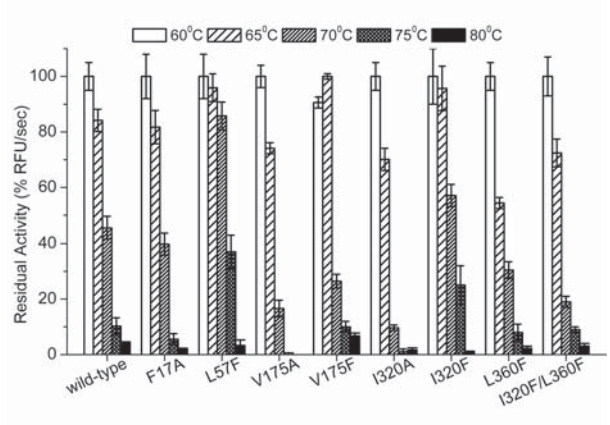

Figure 4.15:
Thermostability of Mutants. Each lipase is incubated at the given temperature for 30 minutes using 4MU-caprylate as substrate. Assays are performed at pH 7.25 and room temperature. Percent activity is shown with error bars representing standard deviations from duplicate measurements. The highest activity for each mutant is set to 100%.

4.3.7 Thermostability of Mutants

The thermostability of each mutant is also profiled to see the effects of the mutations. All of the mutants are found restore at least 50% at the temperature range of 60-65^{o}C (Figure 4.15).

4.3.8 Affinities of C4 and C8 for BTL2 Mutants

PMF profiles for mutant BTL2 are constructed from five pulling SMD trajectories and illustrated in Figure 4.16 A and B for C4 TAG and C8 TAG, respectively.The relative binding free energies of both TAGs for mutant BTL2s are determined at 15 Å of extension as in wild type BTL2, and the values are listed in Table 4.7. The mutations are chosen such that they will reduce the affinity of C8 but wont have an adverse effect on the binding of C4. As seen from Table 4.7, this expectation has mostly borne out; the affinity of C4 has increased for all mutants except for V175F and the affinity of C8 has decreased for all mutants except for I320F. Among the favorable mutations for C4 over C8 selectivity, L360F provides the most promising candidate followed closely by F17A.

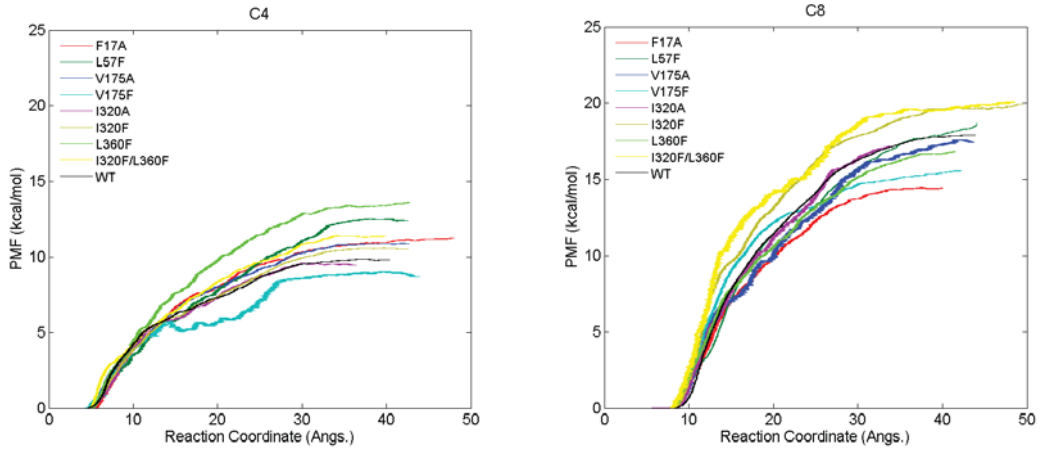

Figure 4.16:
Mutant PMF Profiles. PMF profiles of the mutants, F17A, L57F, V175A, V175F, I320A, I320F, L360F and I320F/L360F are shown for C4 ligands (A), C8 ligand (B) in color series and PMF profile of wild-type is plotted in black color to compared the relative changes in the binding free energy. 5 samplings are used to construct mutant PMF profiles.

Table 4.7:

Binding Free Energy of ES Complex and Specific Activities. [a]The binding free energy are calculated by the implementation of SMD with Jarzynskis Equality. Errors are obtained from 10 pulling trajectories for wild-type and 5 for mutant complexes. [b]Fold-changes in the specific activities of mutants with respect to BTL2 is re-used from Table 4.5.

Lipases	**Tributyrin (C4)**		**Tricaprylin (C8)**	
	ΔG_S^a $(kcal/mol)$	Fold-change[b]	ΔG_S^a $(kcal/mol)$	Fold-change[b]
BTL2	-7.2 $\pm$0.9	1	-15.2 $\pm$1.2	1
F17A	-7.8 $\pm$0.8	0.89	-13.0 $\pm$1.1	3.65
L57F	-7.9 $\pm$0.9	0.63	-14.5 $\pm$1.4	1.08
V175A	-8.0 $\pm$0.4	0.70	-14.0 $\pm$1.0	0.81
V175F	-5.3 $\pm$0.7	2.69	-13.9 $\pm$1.3	3.74
I320A	-7.2 $\pm$0.9	0.50	-15.5 $\pm$1.4	0.81
I320F	-7.4 $\pm$0.4	1.01	-17.2 $\pm$1.2	1.27
L360F	-9.7 $\pm$0.7	1.04	-14.0 $\pm$1.1	3.09

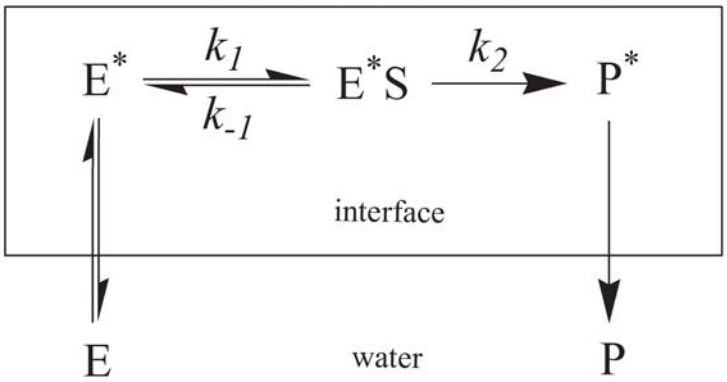

Figure 4.17: Interfacial Kinetics.

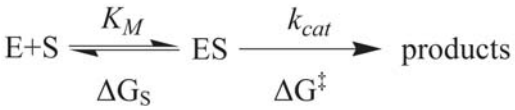

Figure 4.18: Kinetics at the Interface.

4.3.9 Activation Barrier in C4 and C8 Hydrolysis

The experimental determination of lipase kinetics is quite labor-intensive because the physical properties of the lipid substrates might pose extra challenges in the assays. Moreover, accurate determination of the activation barrier for the hydrolysis reactions might not be possible in a straightforward manner due to two sequential interactions of the lipase with the lipid substrates in the interfacial kinetics model (Figure 4.17). Therefore, in this thesis we do not determine the kinetic parameters for wild-type and mutants in the usual way. Instead we aim to obtain a relative estimate of the activation barrier in mutants with respect to wild-type for C4 and C8 TAG. If we consider the activated lipase, the scheme in Figure 4.17 simplifies in to Figure 4.18.

This activation barrier ($\Delta G_T^{\ddagger}$) to be estimated is composed of two terms, the binding energy of ES complex (ΔG_S) and the energy required for transformation of substrate to transition state ($\Delta G^{\ddagger}$).

$$\Delta G_T^{\ddagger} = \Delta G^{\ddagger} + \Delta G_S \tag{4.4}$$

According to equation 4.4, we had to estimate two parameters: the substrate binding affinity (ΔG_S) and the activation energy ($\Delta G^{\ddagger}$) of the hydrolysis. The former (ΔG_S) corresponds to the binding affinity of ES complex whereas the latter ($\Delta G^{\ddagger}$) corresponds to that of $ES^{\ddagger}$ complex. Here, we have used a computational model to estimate the binding affinity of ES complex since it corresponds to the single TAG binding to catalytic cleft of BTL2. The estimation of the binding affinity of ES complex provides us with the binding constant. The binding constant for ES complex is given as follows in Michealis-Menten scheme [198],

$$K_S = k_1/k_{-1} \tag{4.5}$$

But if we consider the revision made by Briggs and Haldane, the equation 4.5 is rewritten as following [199],

$$K_S \approx K_M \quad where \quad k_{-1} \gg k_2 \tag{4.6}$$

In this aspect, we are able to determine the binding free energy of substrate which indeed gives us an estimation of K_S. To predict the k_{cat} constant, which will gives us the estimate of the second term: the activation energy ($\Delta G^{\ddagger}$), we have used two different assays at saturating substrate conditions. Since, we have observed less than 10% of change in the V_0 at two different substrate conditions, it is assumed that the steady state conditions are reached for which the total enzyme concentration is given as,

$$[E_0] = [ES] \tag{4.7}$$

and rate reaction is measure of turnover at saturating substrate conditions,

$$[S] \approx K_M, \quad v = \frac{k_{cat}}{E_0} \tag{4.8}$$

Equation 4.8 gives us the estimate of the activation energy since the activation barrier is given as,

$$\Delta G_T^{\ddagger} = -k_B T \ln(k_{cat}/K_M) \tag{4.9}$$

where k_B is the Boltzmann constant and T is the absolute temperature.

According to ref. [200], deacylation is substantially slower irrespective of the chain length of triacylglycerol substrates. The change of the activation energy of catalysis is typically given by comparing k_{cat}/K_M of mutants with wild-type at conditions where $S << K_M$ so that k_{cat}/K_M is acquired under saturating conditions. Here, we couldn't measure k_{cat}/K_M directly because the emulsion led to inconsistencies. Instead, we calculated K_S and assumed it was K_M by the assumption that $k_2 << k_{-1}$. Also, we calculated k_{cat} directly under saturating conditions, thus we are assuming that there are no allosteric interactions or feedback-related changes of rate.

In TAG hydrolysis, substrate in the ES complex is found in the unaltered form which changes to transition state in the $ES^{\ddagger}$ complex. With respect to this paradigm, at a given substrate concentration, catalytic preference depends on the individual binding free energies of ES and $ES^{\ddagger}$ complexes (ΔG_S and $\Delta G^{\ddagger}$) [191]. We have calculated the former using SMD with Jarzynski's equality and the latter is measured in experiments quantifying specific activities for all mutants. The fold changes in these activity values with respect to wild type BTL2 should correspond to relative contribution of mutant $ES^{\ddagger}$ complexes to $\Delta\Delta G^{\ddagger}$ [191, 192]. From the values given in Table 4.8, we see that variations in specific activities are relatively smaller (<5-fold, < 1 kcal/mol) for both substrates, suggesting negligible contributions from $\Delta\Delta G^{\ddagger}$ to $\Delta\Delta G_T^{\ddagger}$ (Table 4.8).

Table 4.8: Gibbs Diagram Values. Units are kcal/mol for whole table. $^{a}\Delta\Delta G_S$ are calculated by ΔG_S values for mutants with respect to wild-type in Table 4.7. $^{b}\Delta\Delta G^{\ddagger}$ values are calculated using the fold changes in specific activity of mutants with respect to wild-type with an exponential factor. $^{c}\Delta\Delta G_T{}^{\ddagger}$ is calculated from the Equation 4.4.

Lipases	**Tributyrin (C4)**			**Tricaprylin (C8)**		
	$\Delta\Delta G_S^a$	$\Delta\Delta G^{\ddagger}$	$\Delta\Delta G_T^{\ddagger}$	ΔG_S^a	$\Delta\Delta G^{\ddagger}$	$\Delta\Delta G_T^{\ddagger}$
F17A	-0.6	0.1	-0.7	2.2	-0.8	1.4
L57F	-0.7	0.3	-0.4	0.7	0.0	0.7
V175A	-0.8	0.2	-0.6	1.2	0.1	1.3
V175F	1.9	-0.6	1.3	1.3	-0.8	0.5
I320A	0.0	0.4	0.4	-0.3	0.1	-0.2
I320F	-0.2	0.0	-0.2	-2.0	-0.1	-2.1
L360F	-2.5	0.0	-2.5	1.2	-0.7	0.5
I320F/L360F	-1.2	0.2	-1.0	-2.2	0.3	-1.9

In contrast, variations in the binding free energies of mutant ES complexes ($>$ 2 kcal/mol) have greater effect on the catalytic preference (Table 4.8 and Figure 4.19). Thus the relative binding free energy of the ES complex is much more critical than that of ES‡ for assessing catalytic preference of C4 and C8 TAGs, particularly for the BTL2 mutants considered in this study.

4.3.10 Chain-length Selectivity of Mutants

Among the four mutations carried out in the *sn*-1 pocket (V175A, V175F, I320A and I320F), I320A and I320F do not have any significant effects on catalytic preference towards C4 TAG (Figure 4.19), consistent with our strategy that volume adjustments near I320 should not influence C4 affinity. Moreover, I320A mutant does not affect C8 preference which indicates that alanine substitution at this position has a null effect on interactions with TAG. Surprisingly, I320F mutant enhances the catalytic preference towards C8 (negative $\Delta\Delta G_T^{\ddagger}$). It appears that I320F creates a more hydrophobic environment for the acyl chain without restricting the available volume (Figure 4.9), which helps to increase the affinity of C8. Thus, contrary to our expectations, phenylalanine substitution in I320 is

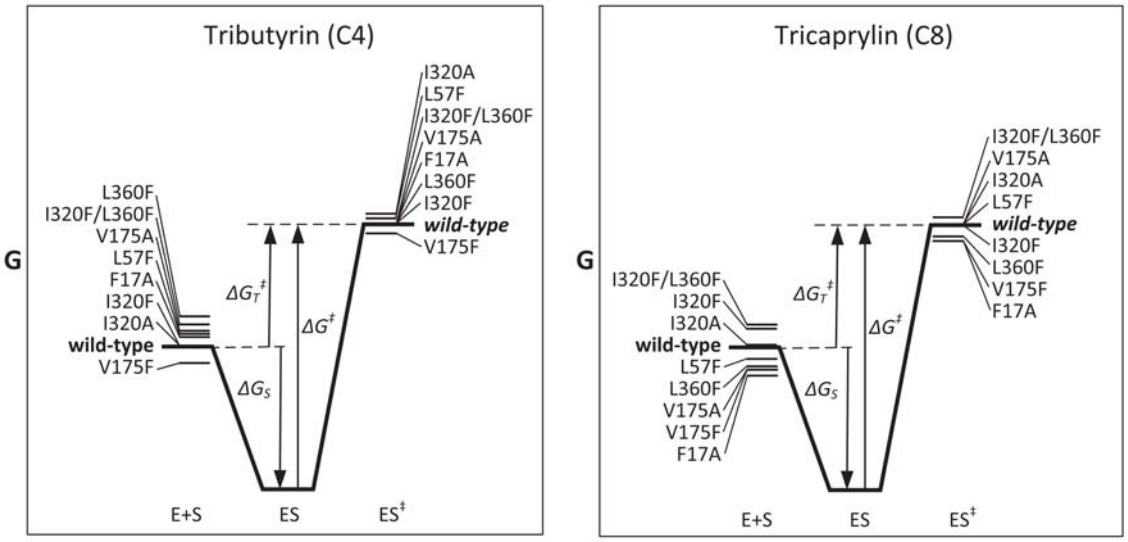

Figure 4.19:
Gibbs Energy Diagrams of Hydrolysis Reactions. Gibbs energy diagram of C4 and C8 TAG hydrolysis for each lipase are given and normalized to the level of ES complex formed by BTL2. Binding free energy changes of ES complex formed by lipases are determined using SMD calculations and the binding free energy changes of transition state complex of the lipases are is calculated by the fold-changes in the specific activities of mutants with respect to wild-type for each substrate (Table 4.4). The catalytic preference of hydrolysis are calculated from Equation 4.4 Table 4.7 is used for all of the values used to rank enzyme, substrate (E+S) and transition state complex (ES) relative to enzyme-substrate complex ($ES^{\ddagger}$).

not effective in reducing the cleft volume for C8.

Another residue in the *sn*-1 pocket, V175, is located in the loop connecting to the N-terminus of α6 helix of lid which is not close to the oxyanion hole (Figure 4.9). The position V175 is investigated in two mutations, V175A and V175F. Comparing the relative affinities of the ES complex formed by V175F (Table 4.6), we observe a weaker binding of both TAGs to the mutant form than BTL2, resulting in a decrease in the catalytic preference towards both TAGs (Figure 4.19). It appears that V175F reduces the available volume for both C4 and C8 ligands, and therefore is not very useful for boosting the selectivity of C4 over C8. For V175A, on the other hand, the binding affinity and catalytic preference are enhanced for C4 but reduced for C8 (Figure 4.19 and Table 4.7). V175A leads to insignificant volume change in the cleft, and thus is not expected to have an effect on C4/C8

selectivity. The fact that it does indicates that V175 is the optimal residue for C8 at this positionchanging it to phenylalanine reduces the volume available to C8 while changing to alanine reduces the hydrophobic interactions. In contrast, V175 appears to be volume limiting for C4 binding, so its mutation to alanine improves the binding affinity. The mutations carried out in the *sn*-3 pocket; F17A and L57F decrease catalytic preference towards C8 and increase it towards C4. For F17A mutation, although a volume increase in the *sn*-3 pocket is presumed to enlarge cleft volume, the preference towards C8 chains is reduced while it is increased for C4 (Figure 4.9) indicating that F17A actually reduces the stabilization of C8 while slightly increasing it for C4. The chain-length selectivity is changed for L57F similarly but not as pronounced as F17A. Compared to F17, L57 is positioned further from the C8 chains (Figure 4.9), and this explains the less pronounced effect of L57F on C8 binding. Due to its position in the *sn*-2 pocket, L360 is not considered to be a critical residue for the catalytic machinery, yet L360F increases the catalytic preference towards C4 the most while reducing it towards C8 (Figure 4.9 and Table 4.7). The finding for C8 is consistent with our mutagenesis strategy of restricting the volume available to C8 but the increase in the binding affinity of C4 is unexpected. Clearly phenylalanine provides a more favorable hydrophobic environment for C4 than valine does. The net result is that L360F provides the most favorable mutation in BTL2 for boosting the enzyme selectivity for C4 over C8.

5 IMPACTS OF THE LID TRYPTOPHANS ON THE ACTIVITY, STABILITY AND AGGREGATION

5.1 Summary

Here, the lipase originating from *Bacillus thermocatenulatus* (BTL2) is a thermoalkalophilic enzyme that exhibits aggregation behavior and is stable at high temperatures [201]. Aggregation might be a strategy for thermostability and this process might involve hydrophobic interactions as a result of high hydrophobic amino acid content (33.2%) in BTL2. Particularly, the lid region in BTL2 are rich in aromatic residues which might be involved in these hydrophobic interactions that may be critical to the aggregation. In this sense, we find two conserved tryptophan (W212 and W235) in the BTL2. Previously, tryptophans are found in many regulatory functions in the lipase structure, especially the ones that are found in the lid region. From this perspective, these tryptophans in the lid region of BTL2 are of particular importance in this chapter. Via protein engineering approaches we have obtained three recombinant lipases (wild-type, W212A and W235A mutants) and characterized all of them in a series of assays for thermostability, thermoactivity and aggregation. Molecular dynamics is also used to see the molecular impact of W212A on the dimer stability of BTL2.

5.2 Methods

5.2.1 Mutagenesis, Expression and Purification

Cloning strategies, production and purifications are explained in Chapter 4.2. Here, the procedures are performed similarly with the only modification in the site-directed mutagenesis step: The mutagenesis with OE-PCR is carried out using the mutant primers pairs given in Table 5.1.

Table 5.1: Mutant Primers for W212A and W235A.

Mutation	Direction, 5-3	Sequence, Mutant Codon, in bold
W212A	forward	CGATTTTAAGCTCGACCAA**GCG**GGGCTGCGCCGCCAGC
	reverse	GCTGGCGGCGCAGCCC**CGC**TTGGTCGAGCTTAAAATCG
W235A	forward	GGCTCAAACGATCCCCTGTT**GCG**ACGTCGACGGATACTG
	reverse	CAGTATCCGTCGACGT**CGC**AACAGGGGATCGTTTGAGCC

5.2.2 Lipase Assays

Titrimetric lipase assays are performed to measure rate of reactions using C4 and C8 TAGs as substrates at 150 mM in the presence of 5% gum arabic. The reaction conditions given in Chapter 4.2. For fluorescence assays, lipase activity is measured in a 96-well black microtiter plate using 4MU-caprylate as substrates in 100 mM Tris-Cl at pH 7.25. Gemini XS (Molecular Devices) is used to measure 4MU-fluorescence using an excitation wavelength of 355 nm and an emission wavelength of 460 nm in a kinetic fashion. The measurements are made at every minute for 1 hour for all analysis. All measurements are performed in duplicates and initial velocities are calculated using SoftMaxPro Software (Molecular Devices).

5.2.3 Thermostability and Thermoactivity Assays

Purified lipases are used in fluorescent assays with 4MU-C8 to determine the optimum temperature range for their stability and activity. Thermostability is profiled by quantifying the residual activity of lipases after incubation for 30 minutes at the given temperatures. In the temperature activity assays, eight measurements (including time zero) in 7 minutes are carried out for each temperature and fluorescence produced at each interval (1 min) is captured by the termination of enzymatic activity in diluted HCl (10 mM). After all measurements from each temperature are completed, end-point assays is performed in Gemini XS fluorometer at high sensitivity.

5.2.4 ANS Fluorescence

Fluorometry is performed by the described procedure in the ref. [202]. In this procedure 1-anilinonaphthalene-8-sulfonate (ANS) is used as the fluorescent probe and the fluorescent readings are conducted via Gemini XS fluorometer at the excitation wavelength of 360 nm and the emission of 460 nm. Three lipases are added to the final concentration of 50 μM in

10 mM Tris-Cl at pH 7.25. ANS is supplied at 50 μM in the mixture. The protein solutions are incubated at the corresponding temperatures without ANS and then left for 5 min in the presence of ANS before readings. The fluorescence is normalized to the background ANS fluorescence collected from samples without protein.

5.2.5 CD Spectra

CD spectra are collected for samples, BTL2, W212A and W235A using J-815 Circular Dichroism Spectropolarimeter (Jasco) in N_2 atmosphere. CD spectra is collected in the 190-350 nm with 1.0 mm path length cuvettes containing maximum 10 μM of lipase. Three tracings are used to collect final spectra for far-UV which is corrected for the background. Mean residue ellipticity $[\theta]$ is calculated from the equation below,

$$[\theta] = 100 \times \theta \times M/(C \times l \times n) \tag{5.1}$$

100 is the correction factor for the cell length which is given as l in centimeters, M is the molecular mass which 43 kDa for BTL2, n is the number of residues which is 389 for BTL2 and C is the concentration in mg/ml.

5.2.6 Dynamic Light Scattering

Dynamics light scattering (DLS) is conducted using Zetasizer Nano XS (Malvern Instruments). Filtered samples are contained 10 μM lipases in 10 mM Tris-Cl at pH 7.8. The data analyses are performed using corresponding software supplied. DLS data is used to analyze the change in the aggregation behavior of BTL2 and its mutants but not to determine molecular weight of the samples. The hydrodynamic radius (R_H) of samples and the standard deviation of the radius (C_P) are collected from the mean value of the monomodal peaks and from the width of the those peaks, respectively. C_P/R_H ratio is used as the polydispersity index (PDI) which is a measure of aggregation. Lower values than 15% in PDI

is used as the cutoff value for monodisperse samples [203]. The statistical parameters generated by the analysis software are interpreted by the analysis software (Zetasizer Nano). DLS data obtained from the cumulant analysis are shown on correlograms for the quality of the measurements and samples. The correlation function and cumulants fits data found in the datalog, are used to assess the quality of the results. The completeness of the correlation function and the goodness of the fit of each measurement are shown in the Appendix A.2 for each measurement.

5.2.7 Zinc Tolerance

The crystal structure of this lipase has two different metal ions (zinc and calcium) thus the effects of zinc on wild-type and W212A activity is investigated and the percentage hydrolysis of 4MU-C8 substrate is calculated. The lipases are supplied at 10 nM in assays and the chloride salts of calcium and zinc are supplied in a concentration gradient from 0.6 mM to 5 mM.

5.2.8 Molecular Dynamics and *in-silico* Mutagenesis

To understand the molecular basis of W212A mutation in dimerization, we have performed molecular dynamics (MD) with closed structure of *Bacillus stearothermophilus* L1 lipase (PDB ID: 1KU0) [156] which has 98% of sequence homology with BTL2[1]. This structure is found in the dimer form composed of two lipase molecules as crystallographic asymmetric units. The dimer structures are used in MD studies in the absence of external forces. The dimer systems are composed of 60,000 atoms placed in a water box with dimensions 73x82x91 $Å^3$. 4 different systems are generated: 2 for the wild-type dimer with or without Zn and 2 for the W212A mutation with or without Zn. The dimer has W212A mutation in chain B for which W212 is found at the dimer interface both chains. The systems are

[1]The sequence alignment is given in Figure 3.2

generated after solvation with a 100 mM NaCl solution. We have performed MD simulations for the resulting systems using the NAMD program [130] with the CHARMM22 parameters [204,205] which include the CMAP correction [206,207]. An NpT ensemble is used in MD simulations with periodic boundary conditions, and the long-range Coulomb interactions are computed using the particle-mesh Ewald algorithm. Pressure is kept at 1 atm and temperature at 300K using the Langevin pressure and temperature coupling. A time step of 2 fs is used in all MD simulations. The trajectory files are obtained in every 2000 fs. Following a 2000 step energy minimization, the dimer systems are simulated for 10 ns. In both systems, root mean square displacements are measured for the backbone atoms C, N, CA.

5.3 Results and Discussion

5.3.1 Structural Investigation of W212 and W235

The tryptophan found in the lid region of BTL2 are investigated using the open (PDB ID: 2w22) [33] and closed conformation (PDB ID: 1KU0) [156] of lipase structures. These two structures of BTL2 and L1 lipase share 95% sequence homology. These investigations might reveal useful information at the atomistic level. Figure 5.1 gives the open and closed structures. The lid region, active site and two tryptophan are shown in Figure 4.6 to see the differences in W212 and W235 residues with respect to open and closed states. Although the contacting residues for W212 are increased in the open state, the bulky side chain in W212 shows the same conformation in the both states (Figure 5.1). However, it is not the case for W235. The side-chain of this tryptophan is not rigid such that it is exposed to solvent in the open state whereas it is buried in the closed state. According to the Figure 5.1, the activation BTL2 includes the movement of the bulky side chain of W235 and owing to its exposed side chain in the open state, W235 is a candidate to mediate the hydrophobic interactions taking place at the substrate interface. In this regard, we start with the activity

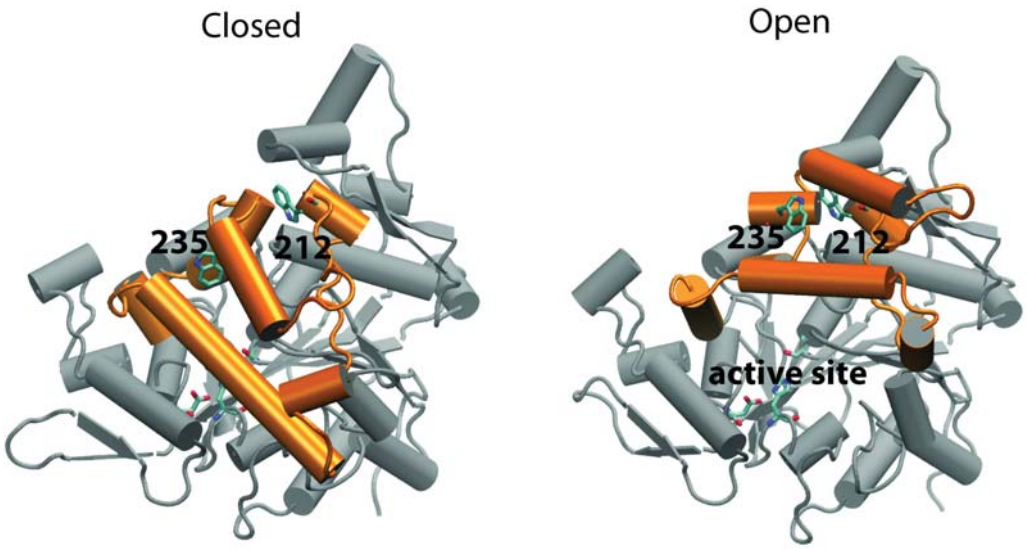

Figure 5.1:
Lid Tryptophans in the Closed and Open Forms. Closed (PDB ID: 1KU0) and open (PDB ID: 2W22) states are shown in cartoon and lid region is colored in orange for both. The lid tryptophan and the active site residues are labeled.

assays to observe impact of the W235A mutant on the lipase activity and chain-length selectivity.

On the other hand, W212 does not change its conformation in the open state, in fact W212 side chain is found buried in both of the crystal structures 1KU0 and 2W22. In the former, W212 is found in the dimer interface (Figure 5.2) and in the latter it is found covered by one of the lid helices (Figure 5.2). Conversely, W212 is found buried in the open state whereas it is partially buried in the closed state. Unlike to W235, it seems that the bulky side chain of W212 is not available for the hydrophobic interactions taking place at the interface to modify the chain-length selectivity. Nevertheless W212 might form other interactions that are required for the overall stability of the lipase since it decreased the overall lipase activity without affecting the chain-length selectivity. W212A has its side chain partially buried in the closed state whereas in the open state, W212 is buried under the short lid helix (α6). In this sense, W212 might be contributing to the stability of the open conformation. In the same sense, the closed structure [156] crystallized in the

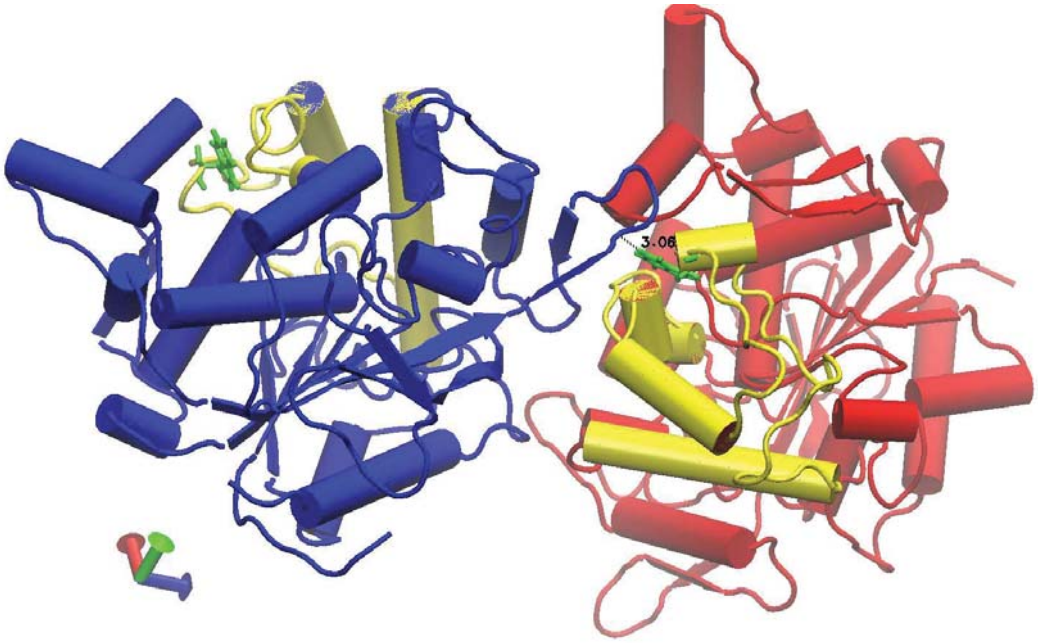

Figure 5.2:
W212 at the Dimer Interface. 1KU0 structure is shown in cartoons revealing secondary structures for chain A (blue) and chain B (red). Yellow is used to color the lid region (residues 175 to 240) for both. The tryptophan in chain B is 3 Å apart from G275 in chain A.

dimer form at the interface that W212 in chain B is found in the close proximity to other chain (Figure 5.2). Similar to the open state (2W22), W212 is found partially buried by the intra-molecular interactions in the dimer. These structural investigations of the closed and open conformations as well as dimer forms of this lipase support the idea that W212 might still be critical to other macroscopic properties of BTL2 such as thermostability and aggregation [201].

5.3.2 Cloning, Expression and Purification

The analytic agarose gel results are given in Figure 5.3 for LIC of W212A and W235A. Two selected positive clones are sequenced in both directions. The mutations are confirmed similarly by BLAST. Expressions are confirmed in LB-Rhodamine plates and mutant lipases are visualized in SDS-PAGE and zymogram (Figure 5.4). These results confirm the

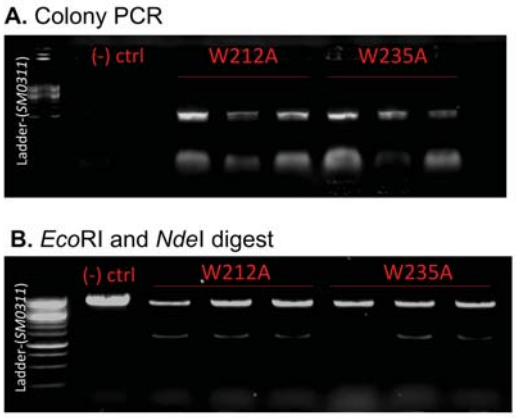

Figure 5.3:
Cloning Confirmations. (A) Colony-PCR reactions using three colonies for W212A and W235A transformants. Negative control lanes given as (-) ctrl are performed to test carry-over contamination from primers. First lane is for forward primer and second is for reverse. The expected size for both mutants is 1200 kb. (B) 3 recombinant plasmids from each mutant are co-digested by *Eco*RI and *Nde*I. The positive clones should give two fragments at 5900 kb -backbone- and 1300 kb -insert-. (-) ctrl is empty vector. The first two positive colonies are selected for sequencing. The electrophoresis is performed at 100 V for 30 minutes using TBE buffer. DNA-Ruler 1 kb ladder (SM0311 Fermentas, see Appendix A.1) is used for molecular weight references.

functional expression of these mutants in this heterologous host. Subsequently lipases are purified (Figure 5.5) following similar techniques explained in Chapter 4.2.

5.3.3 Chain-length Selectivity

Triacylglycerols are the natural substrates for lipases. BTL2 has shown to have the highest activity towards tributyrin (C4) and half of that activity towards tricaprylin (C8) [162]. The rate of hydrolysis is measured for all lipases by titrimetry using emulsified tributyrin (C4) and tricaprylin (C8) substrates at 55^oC. The results are given in Figure 5.6 and suggest that both of the trp mutants decrease C4 activity compared to wild-type BTL2. Tricaprylin activity, on the other hand, is not affected by W235A but is decreased by W212A. Considering W235A mutation, it seems that this tryptophan is required for efficient hydrolysis of tributyrin whereas W212A has negative impacts on the overall lipase activity regardless of the chain-length of the substrate suggesting that W212A might have affected the overall

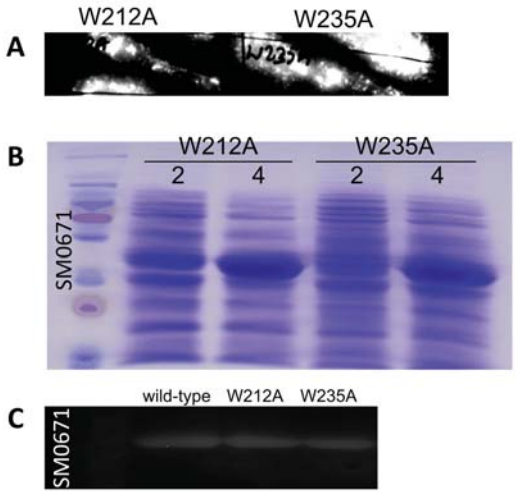

Figure 5.4:
Expression Analyses. (A) W212A and W235A containing SHuffle E. coli cells are plated on LB-rhodamine and the lipase expression is monitored under UV. (B) SDS-PAGE result of W212A ad W235A expressions. Fermentas SM0671 (see Appendix A.1) is used for molecular weight determination. Soluble fractions are obtained from samples of 2 and 4 hours of expression. (C) Zymography analysis is performed using 4-MU-butyrate. The samples obtained from 4^{th} hour are used for large volume expressions of W212A and W235A.

stability of the lipase at 55^oC but not chain-length selectivity of the lipase. According to the comparative analysis of the closed and open lipase structures (Figure 5.6), W235 is considered to be a candidate residue for interfacial binding of BTL2 to substrate interface. Together with activity analysis, we infer that W235 is a critical position for the interfacial binding of tributyrin but not for tricaprylin and alanine substitution at this position might have resulted in the loss of the hydrophobic interactions at the tributyrin interface and thus in the decreased activity. Similar to W235A, *Humicola lanuginosa* lipase contain a trypto-phan in its lid region and mutation of W235 has similarly resulted in a substantial loss of activity towards tributyrin [200]. This indicates that W235 is essential for the distinct C4 selectivity displayed by BTL2.

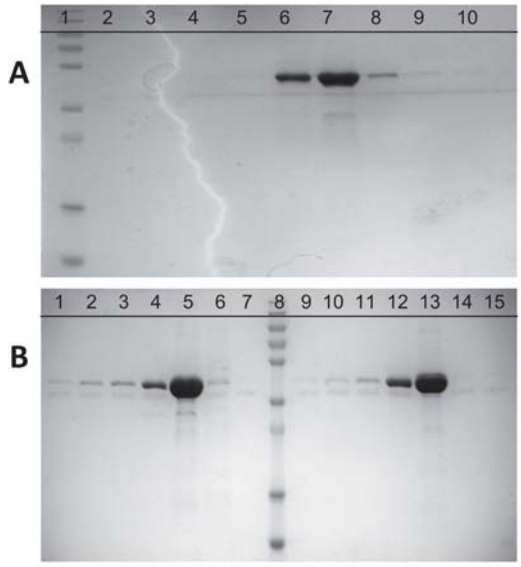

Figure 5.5:
Purification Results. (A) SDS-PAGE result of purification of wild-type lipase. 5 ml-HisTrap column is used for purification and elution is performed in 20 ml by collecting 0.75 ml fractions. Lane 1 is for molecular weight marker (SM0671). First 9 fractions are loaded in 2-10 lanes of SDS-PAGE gel. (B) Purification of W212A and W235A mutant lipases are shown in SDS-PAGE analysis. 1 ml-HisTrap columns are used for purification. Elution is performed in 5 ml by collecting ten of 0.5 ml of fractions. First seven fractions are used in SDS-PAGE. Lane 8: molecular weight marker (SM0671, Appendix A.1), lane 1-7: fractions of W212A and lane 9-15: fractions of W235A.

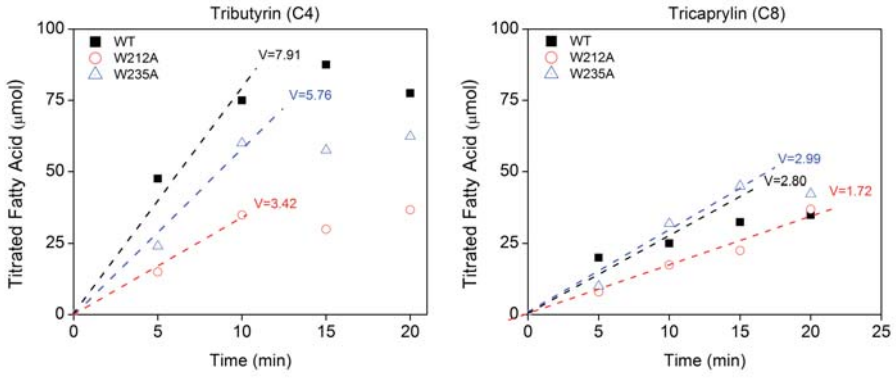

Figure 5.6:
Titrimetric Assay Results. The lipase activity is shown as the titrated fatty acid amount versus time for BTL2 (black), W212A (red) and W235A (blue) using C4 and C8 substrates. The reaction rate (v) is calculated by the slope of the linear fitting functions. The cut-off value for R^2 is set to 0.97 and at least three time points are used for linear fitting.

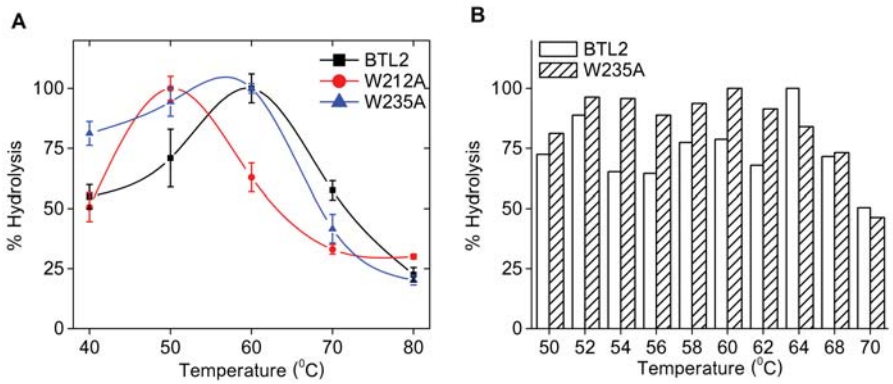

Figure 5.7:
Temperature Activity. The highest activity is set to 100% and percent hydrolysis of 4MU-caprylate is plotted against the reaction temperature. To find out T_{opt} of wild-type, W212A and W235A mutant lipases, temperature range of (A) 40-80^{o}C is scanned by 10^{o}C steps. Error bars show the results of two experiments. (B) 50-70^{o}C is scanned by 5^{o}C steps for BTL2 and W235A.

5.3.4 Temperature Activity Assays

Temperature activity assays are performed to determine the optimum temperature (T_{opt}) for lipase activity. Figure 5.7 A demonstrates that T_{opt} is 50^{o}C for W212A and 60^{o}C for BTL2 and W235A. Further investigation of the 50-70^{o}C range (Figure 5.7 B) for BTL2 and W235A reveals that BTL2 has a T_{opt} of 64^{o}C whereas W235A has showed highest activity at 60^{o}C. Moreover, W235A is more tolerant than BTL2 at the range of 50-62^{o}C. Not shown in Figure 5.7 B but W212A has showed its highest activity at 50^{o}C, in agreement with its profile given in Figure 5.7 A. These findings suggest that W235A mutation has slightly affected the thermoactivity whereas W212A has decreased T_{opt} about 10^{o}C. Among two trp mutants W212A is much more critical than W235A. Although temperature activity assays are performed with fluorescent substrates the overall findings correlate with the finding that W212A mutation is likely to result in the destabilization of the lipase structure at high temperatures.

5.3.5 Effects of Temperature on the ANS Binding

Three lipases are used at 50 μM in fluorometry at different temperatures using ANS as the fluorescence probe for aggregation. Figure 5.9 shows the fluorescence spectra of ANS at 460 nm in the presence of only lipase (white squares) and lipase+5% 2-propanol (black squares). 2-propanol is known to prevent BTL2 from aggregation in aqueous media [203]. In that study, concentrated BTL2 has been shown to be in monomer form in the presence of 5% of 2-propanol. Relying on this, 5% 2-proponal is used as the inhibitor of the aggregation of BTL2 in ANS binding assay. ANS binds to hydrophobic regions in protein surface with high affinity and its fluorescence intensity is significantly increased upon binding to low polarity regions in a protein surface [208]. This makes ANS particularly suitable for the estimation of the surface properties of proteins [209, 210]. ANS and protein-protein interactions -as in aggregation- compete for the similar hydrophobic regions in the surface of BTL2 thus ANS can also be used to probe aggregation [211–213]. According to ANS spectra, the aggregation behavior of BTL2 is clearly dependent on the temperature. At increased temperatures like 65^oC and 75^oC, ANS binding increases which is an indication of deaggregation (Figure 5.9). Additionally, BTL2 activity which is proportional to the available active sites - is significantly enhanced in the presence of 1% cholate[2] at low temperatures whereas activity is increased slightly at high temperatures. Both from the reference and the ANS spectra of BTL2 as a function of temperature agree that BTL2 molecules tend to aggregate at low temperatures. This inference can also be made from the reference that investigates the aggregation and activity relationship of BTL2 [201].

Although it is not possible to comment on the oligomeric form of BTL2, whether it is composed of monomers or dimers, only by looking at the ANS spectra, increased ANS fluorescence reflects an increased binding surface for ANS and thus could be used to compare the BTL2 and mutants by means of their available surface for ANS binding and thus

[2]Cholate is a bile salt and known to inhibit aggregation of BTL2, similar to 2-proponal

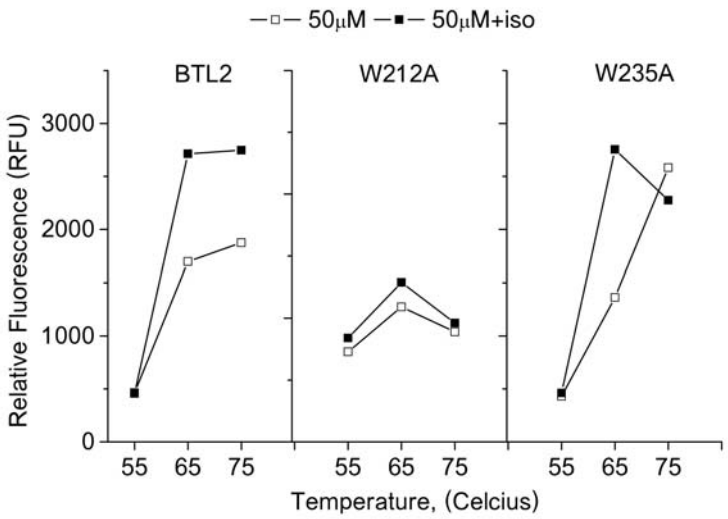

Figure 5.8:
ANS Fluorescence Spectra. ANS binding is shown as a function of temperature for three lipases. The plots are generated using mean values of duplicate assays. White squares corresponds to measurements of only proteins and black squares correspond to the measurements of proteins in the presence of 2-proponal as an inhibitory reagent for aggregation of BTL2 [203].

their aggregation tendency. As can be seen in Figure 5.9, binding of ANS to the BTL2 is higher (2-folds) in the presence of 2-proponal at both temperatures 65^oC and 75^oC. This substantial increase in ANS binding in the presence of 2-proponal indicate that the surface hydrophobicity of BTL2 molecules is shielded by aggregation. Similarly, ANS binding to W235A is increased at 60^oC in presence of 2-proponal but it shows a decrease at 75^oC in the presence of 2-proponal, unlike to BTL2. This might be a result of thermal denaturation of this mutant below 75^oC. Interestingly, W235A shows a decrease in the fluorescence when treated with 2-proponal which is known to inhibit aggregation of BTL2. But aggregated W235A does not show any decrease in the fluorescence suggesting that aggregation may be retarding the thermal denaturation in this mutant. Lastly, W212A has resulted in undifferential ANS spectra at different temperatures or in the presence of 2-proponal. This unaltered ANS binding can be explained by the little tendency of this mutant to form aggregates since 2-proponal has no effects on the fluorescence intensity in W212A. More-

over, at 55^oC ANS binding in W212A is one fold higher than those of BTL2 and W235A, suggesting ANS binding is higher at 55^oC is not competed by other lipases in solution and aggregation is inhibited by W212A.

5.3.6 CD Spectra

To comment on the overall secondary structure of the trp mutants with respect to wild-type, the far UV CD spectra are collected using 10 μM of lipases. Far-UV CD spectra in the 190-250 nm corresponds to the absorption of amide groups and will be used to report the secondary structure of proteins [214]. Typically, all α-helical proteins give two bands of similar magnitude at 222 nm and 208 nm which is accompanied by a positive band around 190 nm. All β-sheet-sheet proteins generally show a negative band between 210-220 nm and a positive one between 195-200 nm nevertheless CD spectra for β-sheet proteins may vary more than that of α-helical proteins due to diverse β-sheet configurations in proteins such as parallel, antiparallel or mixed [214]. Keeping these in mind, the far-UV CD spectra of BTL2 (black), W212A (red) and W235 (blue) are shown in Figure 5.8. Ttwo characteristics of a-helical structure is faint but visible around 208 and 222 nm for BTL2 and W235A. These two lipases also yield a positive band at 190 nm which confirms the presence of helical content for the lipase structure. With respect to β-sheet characteristics, the given CD spectra in Figure 5.8 is no different from each other. The secondary structure of W235A is similar to the wild-type yet W212A appears to be different from wild-type spectra suggesting an instable structure for this mutant. Further investigations are conducted to elucidate the role of W212 for the stabilization of the overall structure.

5.3.7 Effects of Aggregation on Temperature Stability

The thermostability assays are performed by incubating lipases at two different concentrations. Since aggregation is positively correlated with the concentration, we expect to see

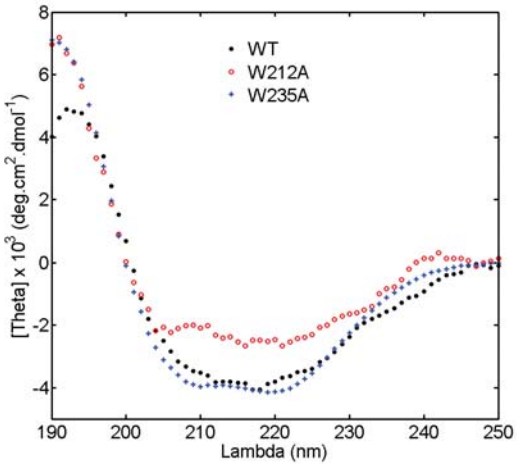

Figure 5.9:
Far-UV CD Spectra. BTL2 (black), W212A (red) and W235 (blue) symbols corresponds to mean residue ellipticity $[\theta]$ as a function of wavelength, λ. The spectra of each lipase is subtracted from blank spectra collected for the buffer system. Data steps for each measurement is 1 nm.

the effects of the aggregation on the thermal stability in these assays. The concentration of BTL2 is increased from 1 μM to 50 μM during thermal incubation to test the effect of aggregation on the thermostability. The results are shown for all lipases in Figure 5.10. BTL2 and W235A restore at least 75% of total activity at 65^oC at 50 μM but this tolerance is almost absent for both when they are incubated at 1 μM concentrations (Figure 5.10). Clearly, in the temperature range of 60-65^oC, thermostability is restored with the increased concentration of BTL2 and W235A. In other words, aggregation of this lipase resists thermal denaturation and contributes to the thermostability of the wild-type lipase and W235A mutation does not affect the positive correlation of the aggregation and thermostability in BTL2. However, the other mutant, W212A does not yield different thermostability profiles at different concentrations. In other words, high concentration of W212A is failed to restore the activity compared to low concentration. This observation together with the decreased thermostability of W212A may in part explaining the decreased stability of W212A mutant by the defect in the aggregation behavior. In the absence of aggregation, W212A is likely to be more susceptible to thermal denaturation. Particularly, this discrepancy in

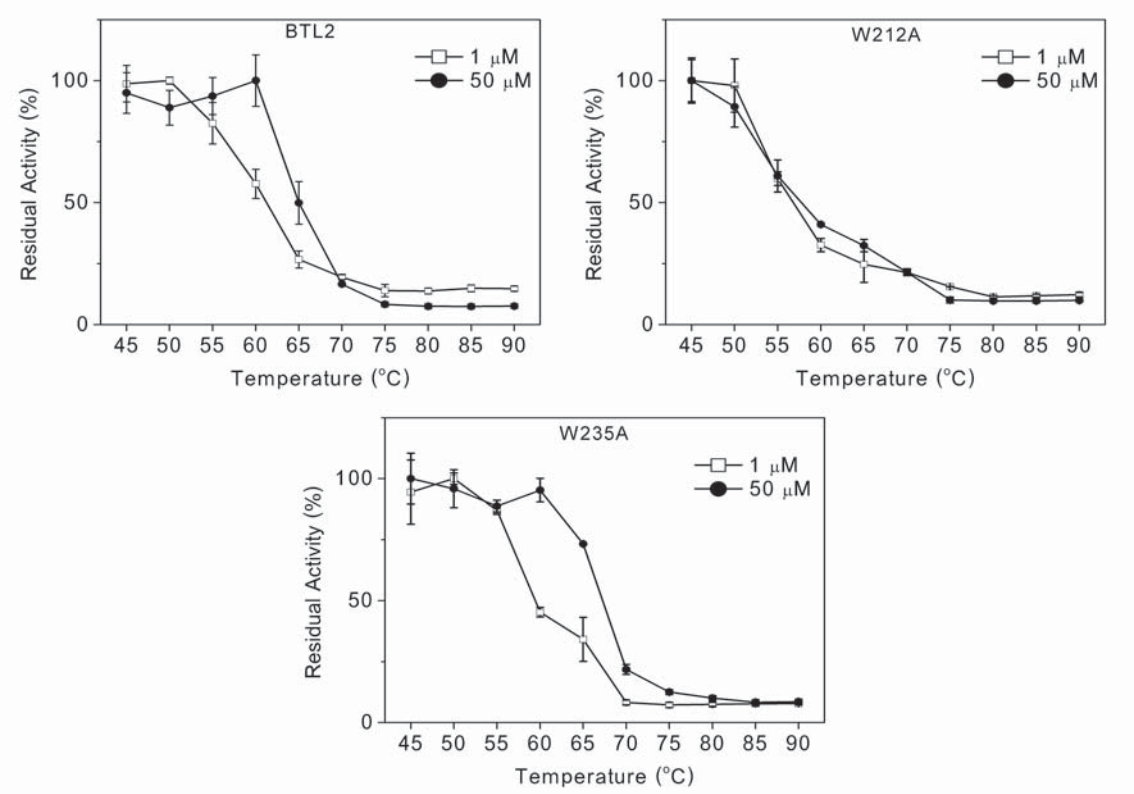

Figure 5.10:
Effects of Aggregation on Thermostability. The lipases are incubated for 30 minutes at the corresponding temperatures and the assays are performed after serial dilutions to the incubated samples such that all lipase assays contain 10 nm of each lipase, 200 μM of 4-MU-caprylate and 0.1 M Tris-Cl at pH 8.0. Residual activity is reported as the percentage of 45^{o}C results. Two different concentration (1 μM, and 50 μM) is plotted. Error bars represent deviations of the two independent experiments.

the wild-type and W212A has enforced further tests to resolve between aggregates and oligomers in these samples.

5.3.8 Dynamic Light Scattering

To confirm the inhibition of the aggregation behavior of BTL2 in case of W212A mutant, dynamic light scattering (DLS) analyses of both lipases are carried out using 2.5 μM concentration of lipase at 25^oC. DLS is a rapid, non-invasive and reliable technique for determination of the globular protein size and aggregation behavior by monitoring the scattering intensity fluctuations and the correlations in the signal decay intensity [215–218]. Since intensity fluctuations are consequences of particle motion which are measured properly in the correlation analysis of the distribution of diffusion coefficients, this technique provides rapid access to size information for the characterization of proteins and it is also very sensitive to the presence of aggregates [218, 219]. Previously in a DLS analysis, the hydrodynamic radius of BTL2 was found to be around 3-4 nm in its monomer form, previously [203]. Figure 4.12. shows the intensity and volume size distributions of BTL2 and W212A obtained from three measurements in this study. All of the distributions are monomodal which is an indication of polydispersity. The peaks that have the largest area are chosen to comment on the majority of the particles in the samples and the results of three measurements are given in Table 5.2. The largest peak size (Rh) in BTL2 is around 72 nm in average, much larger than a monomer (3-4 nm). The width of this peak (C_P) is relatively larger and thus the polydispersity index ($\%C_P/R_H$) is higher in BTL2. In other words, majority of the particles (80%) found in BTL2 are polydisperse, i.e. aggregated. But according to Table 5.2, W212A sample is dominated by the particles with the average size of 3.89 nm, in agreement with the previous finding [203]. Moreover, the percentage of these monomers in W212A is much higher (99%) than other particles. The width of these peaks are very low which also shows the monodispersity with a PDI of $< 10\%$. These

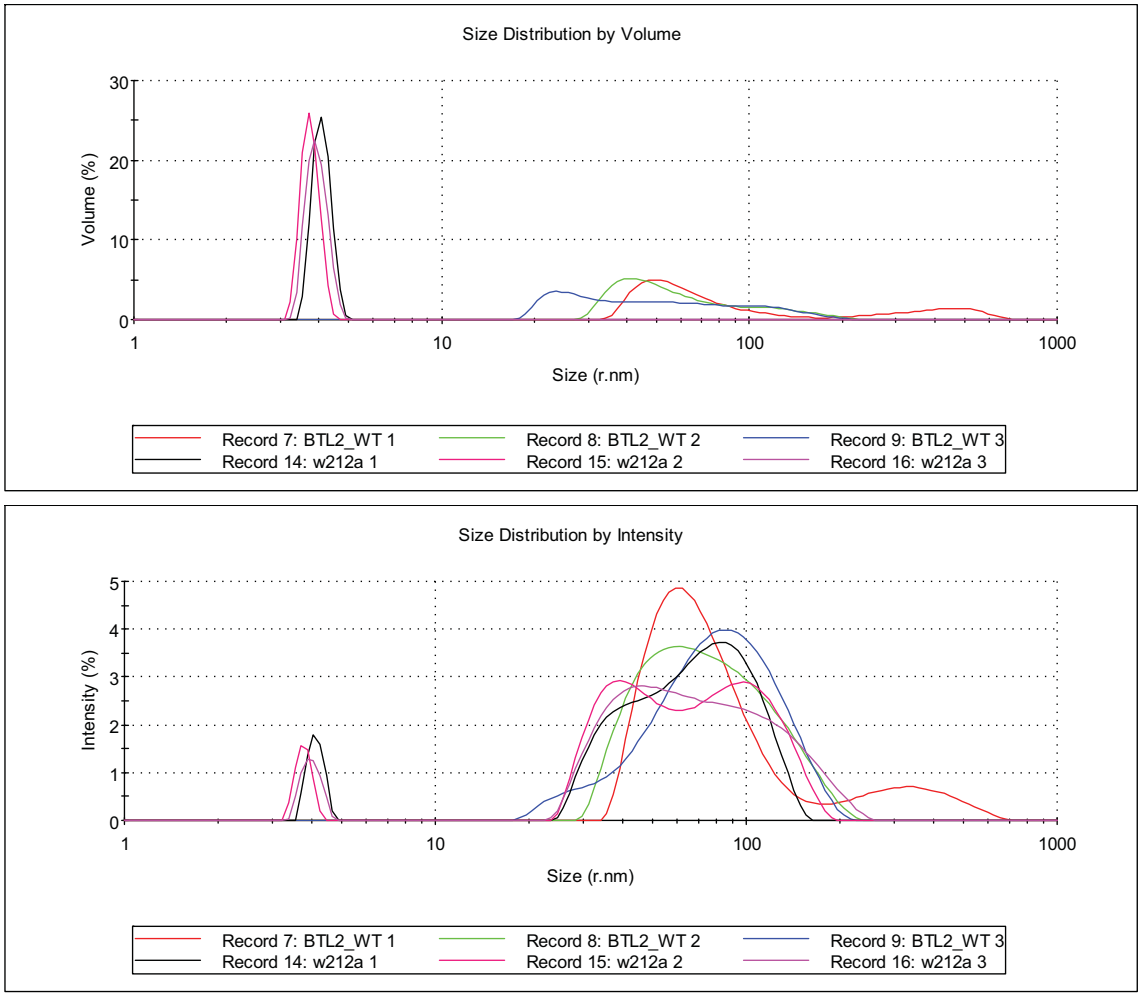

Figure 5.11: Intensity and Volume Distributions of DLS Data.

findings confirms that the W212 is required for aggregation of BTL2 lipase.

As mentioned, the results are obtained from multimodal data for each measurements. Nevertheless, single peaks with the largest area chosen for the monomodal analysis, other particulate at different sizes and intensities are visible in the intensity-size distributions in Figure 5.11. Intensity distribution of BTL2 is typical to a polydisperse sample, i.e. relatively large particles at varying size and shape. In the intensity distribution of W212A (Figure 5.11), the largest contribution is made by peaks that are very similar to BTL2, nevertheless the percentage of these particles are very low in W212A samples. At first these peaks may be considered as dirt or contaminating particles in the sample solution but

Table 5.2: DLS Results.

	nm		**%**	
	R_H	C_P	C_P/R_H	V
BTL2	71.65 ± 12.09	20.35 ± 4.74	28.40	79.66
W212A	3.89 ± 0.19	0.27 ± 0.03	6.94	98.83

these peaks are in very similar in size and shape to the ones observed for aggregated BTL2. Additionally, the measured radius for these relatively large particles in W212A is too low i.e. below micron range to be a dust particle. Thus, these peaks may correspond to the similarly aggregated particles in BTL2 which indicates that a very small portion of particles found in W212A mutant lipases are still aggregating even in the presence of a mutation that restricts aggregation for the majority of particles. In this sense, the aggregation behavior of BTL2 must be a complex mechanism which is likely to be dependent on not a single residue but many other. Still W212A is a very effective mutation to suppress aggregation of BTL2.

5.3.9 Effects of Metals on the Activity

Here, we have analyzed the effects of calcium and zinc on the activity of wild-type and W212A mutant. Both of the metals are found in the crystal structure of open and closed conformations [33,156]. Zinc is found to be required for the stability and activity of BTL2 whereas calcium has not significant affects [156]. Divalent cations like zinc zinc is also found to affect the activity of other thermostable lipase through aggregation [220]. In this regard, we show the effect of increasing calcium and zinc concentration on the activity of wild-type and W212A (Figure 5.12). It seems that calcium decreases the activity of BTL2 whereas it is ineffective in case of W212A (Figure 5.12). The negative effects of the zinc on the lipase activity is retarded in W212A mutant. Considering that metals may promote the aggregation of this lipase [220–222], we see that W212A mutant is significantly tolerant to this aggregation. These results indicate that W212A mutant is a deformed in its zinc

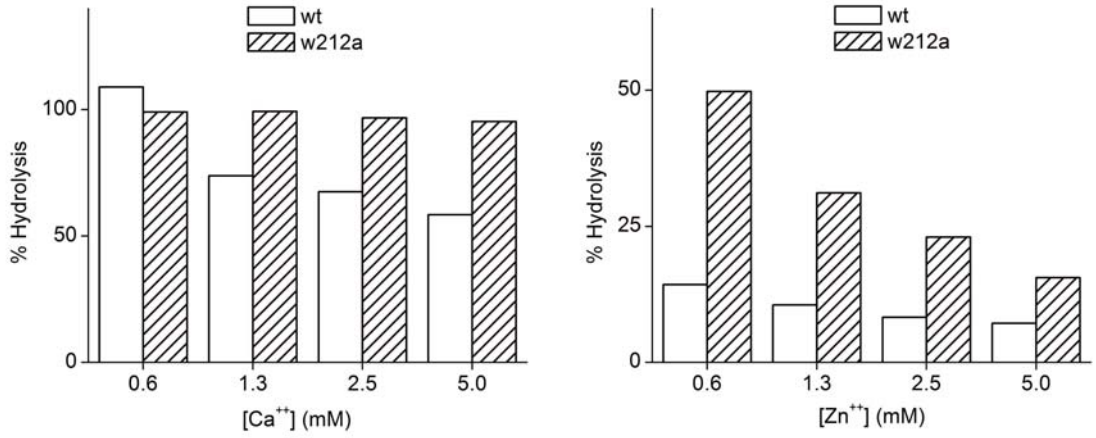

Figure 5.12:
Effects of Calcium and Zinc on Activity. The percent hydrolysis is shown in a concentration gradient of Calcium and Zinc from 0.6 to 5.0 mM. The activity for wild-type and mutant is set to 100% for the samples without Zn. All assays are performed in 0.1M Tris buffer at pH 7.25.

domain so that the zinc does not have significant effects on the activity.

5.3.10 Molecular Dynamics

Dimers with Zinc

To understand the mechanism of W212 in the aggregation of BTL2, we have implemented molecular dynamics (MD) simulations. We have used the crystal structure of the lipase dimer in which both of the monomers are in closed conformation. We thought that the crystal structure that is found in the dimer form (PDB ID: 1KU0) could be used to assess the aggregation of BTL2. We have used 10 ns of simulations for wild-type and W212A mutation and RMSD analysis of the simulations reflect an equilibration of all dimer structures with less than 1 Å (Figure 5.13). For all systems we do not see any movement of the metal ions (2 Ca and 2 Zn) within the structure throughout 10 ns. Upon investigation of the residue at the dimer interface, we have encountered two hydrogen bonds formed by the backbone amide groups of two residues in chain A and the carbonyl oxygen groups of

two complementary residues in chain B (Figure 5.14). The persistence of these hydrogen bonds for all systems is shown in Figure 5.15. Evidently, the most strong pairing (HB I) is the one between L285(NH) of chain A and T1235 (O) of chain B and the relatively weaker one (HB II) is formed between L277 (NH) of chain A and G86 (O) of chain B. Particularly HB I is persistent regardless of W212A mutation Figure 5.15. However, the relatively weaker bond, HB II is destabilized in the case of W212A mutation. According to molecular dynamics findings, we conclude that W212 is required for the stability of one of the bonds at the dimer interface and this partial stabilization of the HB II could explain the loss of aggregation in W212A mutant. W212 is located at the close proximity to the HB II and alanine substitution at this position affects the backbone NH group configuration of L277 residue which leads to dissociation of the HB II in case of this mutant. In this aspect, W212 is a critical residue that hold the dimer together in 1KU0 structure and it is required for aggregation. We have additionally performed molecular dynamics of the dimer systems at 400K to see the effects of the temperature on the structure posed by W212A mutation yet we have observed that at 400K, zinc ion, leaves its coordination shell and moves freely in the wild-type dimer (Figure 5.16). This indeed corresponds to deformation of the tetrahedral zinc coordination which may be a result of parameters for the divalent zinc cation in CHARMM force field which is used here. We have also encountered other cases that also suffered from such deformations in similar zinc coordination sites in molecular dynamics simulations [223]. To solve this movement of zinc, additional restraints may be introduced on the system to prevent the deformation of the tetrahedral zinc coordination. Yet, it was best not to put any restraints in our case since zinc ion is found at the dimer interface and may influence the dimer interactions. In these regards, here we avoid drawing any other conclusions other than Figure 5.16 related to high temperature simulations.

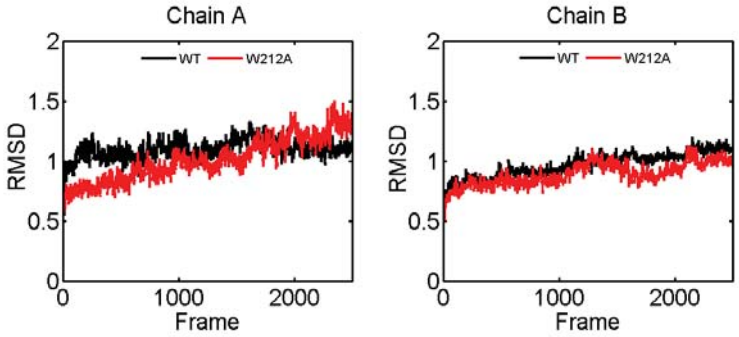

Figure 5.13:
RMSD of the Dimers. The backbone atoms (C, CA and N) are aligned to the minimized structure and RMSD in Angstrom units is shown for wild-type dimer (black line) and W212A mutant dimer for 10 ns simulations. The metals found in the crystal structures (Ca and Zn) are included in these simulations and stayed in their initial configuration without any restraints.

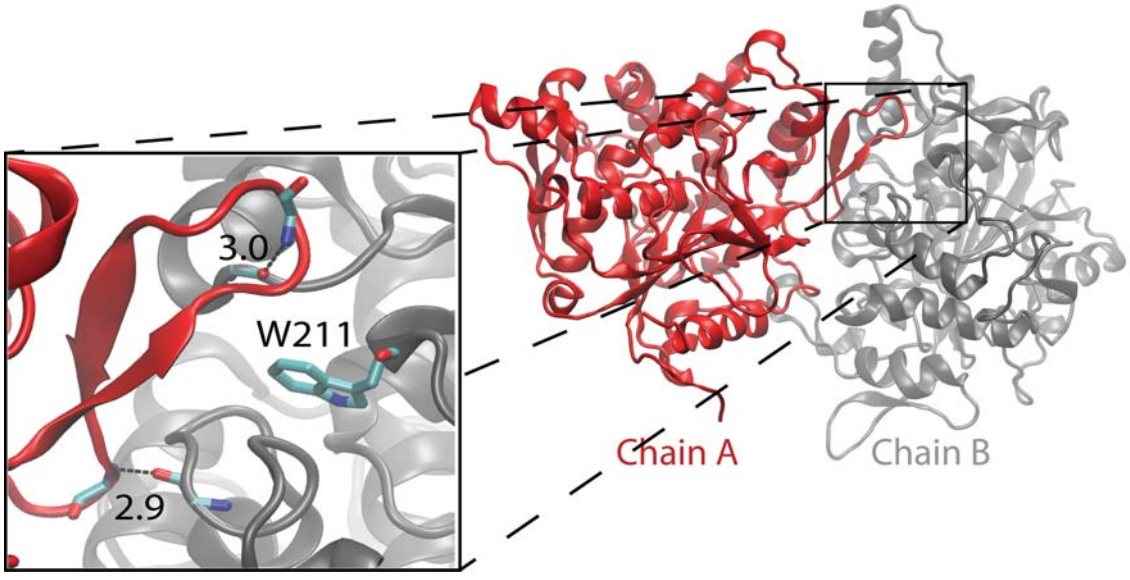

Figure 5.14:
Interactions at the Dimer Interface. Chain A is colored in silver and chain B is colored in tan and two hydrogen bonds between backbone NH (L278) and O (G87) atoms and between NH (L286) and O (T236) are shown.

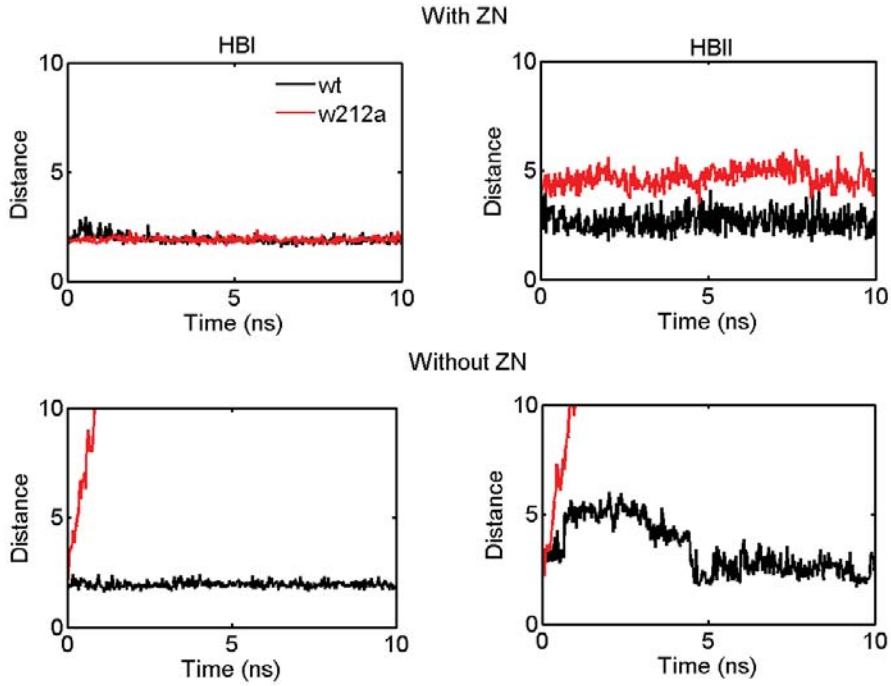

Figure 5.15:
Effects of Zn on Dimerization. (A) The dimer bonds given in Figure 5.14 are measured throughout 10 ns simulations in the presence of both metals. Black line shows wild-type and red line shows W212A mutant dimer. HB I corresponds to the hydrogen bond between L278-NH and G86-O and HB II corresponds to the second hydrogen bond between L285NH-T235-O. (B)Effects of Zn Absence on the Dimerization. Two bonds that are shown in Figure 5.14 are measured similarly for the dimer structures that lacks Zn ions. The simulations of the W212A mutant dimer that lack Zn ions are terminated when a separation of the monomers are visible.

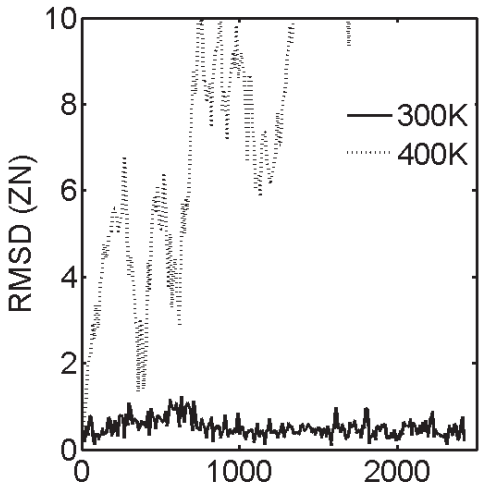

Figure 5.16:
RMSD for Zn at High Temperatures.

Dimers without Zinc

After we see that 400K leads to deformation of the zinc coordination and in turn affects the stability of the wild-type dimer, we decide to remove the zinc ion and see the effects of the zinc absence on the wild-type and mutant dimer. Same hydrogen bonds are measured and given in Figure 5.15. It is obvious that zinc absence is not detrimental alone to the stability of dimer for 10 ns, indeed we see a partial dissociation of the same bond, HB II (Figure 5.15). Actually, zinc absence in the wild-type dimer creates a less drastic effect on HB II than W212A mutation at the interface (Figure 5.15). However, the W212A mutant dimer without Zn breaks both of the hydrogen bonds immediately and results in the dissociation of the dimer structure (Figure 5.15). Yet neither zinc absence nor W212A mutation leads to dissociation of the dimer in MD simulations.

5.3.11 Molecular Impacts of W212 and W235

W235A is particularly important for efficient hydrolysis of tributyrin. This impact of W235A is investigated with respect to the location of its bulky side chain in the closed

and open conformations and W235A is like might be critical for the binding of C4 TAG on the lid and thus for the selectivity of C4. However, we have not found any discrepancy in the thermoactivity and thermostability of W235A. On the other hand, for W212A mutation, we have ended up with differential results than W235A or wild-type. more than we expected. This mutation significantly affects the thermostability of wild-type lipase but we also show that W212A is essential for the aggregation of BTL2. In this sense, decreased thermoactivity and thermostability of W212A is likely to be a result of the loss of aggregation in W212A. In other words, W212A mutation stops aggregating which makes is more susceptible to thermal inactivation. The basis of the impact of the W212A mutation on the aggregation is further investigated in molecular dynamics using a dimer structure of BTL2. The dimer is considered to be a simple model for the oligomerization of BTL2. From MD analysis we show that W212A is particularly critical for the HB II yet we cannot see a complete dissociation of the dimer structure in the case of W212A mutation for 10 ns. Together with this finding that we interpret our MD analysis such that W212A has more acute effects on the one of the hydrogen bonds that hold dimer structure together. However, according to DLS measurements, W212A mutation is shown to yield monomer lipase molecules. It would be safe to conclude as that W212A is essential for oligomerization/aggregation of BTL2 via destabilizing the hydrogens bonds at the interface. Apart from these, we have seen that zinc absence in W212A mutation causes an immediate disruption of both of the hydrogen bonds at the dimer interface. In the experiments to test the effects of zinc in the wild-type and W212A mutation, we found that wild-type lipase activity is inhibited by the minute amount of zinc in the assays however, this inhibition is retarded significantly for W212A mutation. This suggested that zinc promotes the aggregation of this lipase and decreases the available activity site in the assays. Other thermoalkalophilic lipases contain zinc binding domain which is unique to this group among other lipases. W212 is also found at the close proximity to zinc coordinating residues and addi-

tionally, zinc absence was found to affect thermoactivity [156] and thermostability [161] of *B. stearothermophilus* L1 lipase similar to the findings on W212A mutation. The recent studies on other enzymes like cytochrome oxidase [222, 224] was shown to aggregate in presence of divalent cations like zinc and secondary residues in the close proximity to zinc coordination were also found to be critical for the aggregation processes in those enzymes. In this sense, the role of zinc in the aggregation of thermoalkalophilic lipases could be similar to this and W212 is most likely to be important in the aggregation of should also be clarified i.e. whether W212 provides secondary interaction for zinc-induced aggregation. Overall we see that this residue might be essential not only being a lid tryptophan but also for accompanying secondary interactions of zinc domain which could be the trigger of the oligomerization and aggregation this lipase. In these aspects, W212A is a promising mutation that inhibits the aggregation of BTL2 since this lipase aggregates through a very similar mechanism of neurodegeranative protein pathogens in human, like amyloid fibrils and prions [225].

6 CONCLUSIONS

Protein engineering approaches can provide insights for designing novel lipases with improved features. This thesis investigated the selectivity and stability of thermoalkalophilic lipase from *Bacillus thermocatenulatus* via rational design approaches.

Understanding the chain-length selectivity of BTL2 will be useful for applications in industrial sectors using short chain fatty acids because it has a distinct selectivity for short chains. In a combined computational/experimental approach, we have analyzed and modified the chain-length selectivity of BTL2 in terms of catalytic preference. Our results indicate that the L360F mutation will create the most selective enzyme for C4 over C8 followed by F17A and V175A. Further selectivity gains can be achieved by considering multiple mutations, e.g. L360F in combination with F17A or V175A. Because each residue interacts with a different chain, one expects the free energy changes in double mutations to be given by the sum of single mutations.

Our design strategy is simply based on blocking of C8 acyl chains in mutants using in-silico ES complex models of BTL2. The effect of mutations on C8 binding depends on their relative positions in the cleft (i.e. *sn*-1, *sn*-2 and *sn*-3) and also their relative positions with respect to C8 chains. Moreover, mutations are not in direct contact with the oxyanion hole, and hence the specific activity changes for the mutations slightly influence the catalytic preference. ES complex affinities have a much greater influence on the catalytic preference, and these quantities can be estimated from SMD simulations easily. We note that restricting the volume strategy worked only in some cases, while in some other cases, mutations actually increased the affinity of C4, which was unexpected. This shows the complexity of the problem and reinforces the importance of inspecting each case with the same rigor. In

conclusion, we have proposed a novel approach to analyze and alter substrate selectivity of enzymes in terms of relative changes in the binding free energies of ES and ES‡ complexes determined from simulations and experiments. The approaches used in this study could be applied to other lipases to modify their substrate selectivity and produce tailor-made fatty acids.

A practical knowledge on how thermoalkalophilic lipases are stable at high temperatures not only helps in obtaining insights to design novel lipases with desired stabilities but also deepens our understanding of thermostability. Particularly, tryptophans in the lid region of BTL2 are found to be involved in such critical roles. Here, W235A substitution was shown to lower chain-length selectivity i.e. tributyrin/tricaprylin activity whereas W212A had no effects on that. On the other hand, W212A is shown to affect aggregation, thermostability and thermoactivity of wild-type lipase, negatively. Aggregation of industrially important enzymes like BTL2 is a major concern such that aggregation decreases the solubility of the recombinant proteins and thus results in loss of yield of production. Here, W211A mutant represents a promising mutation which is defective in aggregation and thus increases solubility of BTL2. Yet, increased solubility of BTL2 comes with the price in W211A such that it also decreases thermostability and thermoactivity. Industrial conditions may require non-optimal conditions like high temperatures, although decrease in optimum working temperatures of enzymatic processes may aid in lowering the energy consumption of the biocatalysis. Hence, the particular decrease in thermostability of W211A mutation may not necessarily be a disadvantage.

Bibliography

[1] E. Busto, V. Gotor-Fernandez, and V. Gotor, “Hydrolases: catalytically promiscuous enzymes for non-conventional reactions in organic synthesis,” *Chem Soc Rev*, vol. 39, no. 11, pp. 4504–23, 2010.

[2] M. Kapoor and M. Gupta, “Lipase promiscuity and its biochemical applications,” *Process Biochemistry*, vol. 47, pp. 555–569, 2012.

[3] L. G. Frenken, M. R. Egmond, A. M. Batenburg, and C. T. Verrips, “Pseudomonas glumae lipase: increased proteolytic stability by protein engineering,” *Protein Eng*, vol. 6, no. 6, pp. 637–42, 1993.

[4] A. Svendsen, “Lipase protein engineering,” *Biochim Biophys Acta*, vol. 1543, no. 2, pp. 223–238, 2000.

[5] I. Kauffmann and C. Schmidt-Dannert, “Conversion of bacillus thermocatenulatus lipase into an efficient phospholipase with increased activity towards long-chain fatty acyl substrates by directed evolution and rational design,” *Protein Eng*, vol. 14, no. 11, pp. 919–28, 2001.

[6] D. Rotticci, J. C. Rotticci-Mulder, S. Denman, T. Norin, and K. Hult, “Improved enantioselectivity of a lipase by rational protein engineering,” *Chembiochem*, vol. 2, no. 10, pp. 766–70, 2001.

[7] C. C. Akoh, G. C. Lee, and J. F. Shaw, “Protein engineering and applications of candida rugosa lipase isoforms,” *Lipids*, vol. 39, no. 6, pp. 513–26, 2004.

[8] J. Khurana, R. Singh, and J. Kaur, “Engineering of bacillus lipase by directed evolution for enhanced thermal stability: effect of isoleucine to threonine mutation at protein surface,” *Mol Biol Rep*, vol. 38, no. 5, pp. 2919–26, 2011.

[9] H. B. Brundiek, A. S. Evitt, R. Kourist, and U. T. Bornscheuer, “Creation of a lipase highly selective for trans fatty acids by protein engineering,” *Angew Chem Int Ed Engl*, vol. 51, no. 2, pp. 412–4, 2012.

[10] P. K. Sharma, R. Kumar, O. Mohammad, R. Singh, and J. Kaur, “Engineering of a metagenome derived lipase toward thermal tolerance: effect of asparagine to lysine mutation on the protein surface,” *Gene*, vol. 491, no. 2, pp. 264–71, 2012.

[11] M. Fischer and J. Pleiss, “The lipase engineering database: a navigation and analysis tool for protein families,” *Nucleic Acids Res*, vol. 31, no. 1, pp. 319–21, 2003.

[12] A. M. Klibanov, “Enzymatic catalysis in anhydrous organic solvents,” *Trends Biochem Sci*, vol. 14, no. 4, pp. 141–4, 1989.

[13] R. J. Kazlauskas and U. T. Bornscheuer, *Biotransformations with Lipases*, ser. Biotechnology. Weinheim: Wiley-VCH, 1998, vol. 8a.

[14] L. Brady, A. M. Brzozowski, Z. S. Derewenda, E. Dodson, G. Dodson, S. Tolley, J. P. Turkenburg, L. Christiansen, B. Huge-Jensen, L. Norskov, and et al., “A serine protease triad forms the catalytic centre of a triacylglycerol lipase,” *Nature*, vol. 343, no. 6260, pp. 767–70, 1990.

[15] R. N. Farias, M. Torres, and R. Canela, “Spectrophotometric determination of the positional specificity of nonspecific and 1,3-specific lipases,” *Anal Biochem*, vol. 252, no. 1, pp. 186–9, 1997.

[16] L. Vaysse, A. Ly, G. Moulin, and E. Dubreucq, “Chain-length selectivity of various lipases during hydrolysis, esterification and alcoholysis in biphasic aqueous medium,” *Enzyme and Microbial Technology*, vol. 31, no. 5, pp. 648–655, 2002.

[17] T. Godfrey, “Lipases for industrial use,” *Lipid Technol*, vol. 7, pp. 58–61, 1995.

[18] R. Schmid and R. Verger, “Lipases: Interfacial enzymes with attractive applications,” *Angewandte Chemie International Edition*, vol. 37, no. 12, pp. 1608–1633, 1998.

[19] S. Krishna, B. Manohar, S. Divakar, S. Prapulla, and N. Karanth, “Optimization of isoamyl acetate production by using immobilized lipase from mucor miehei by response surface methodology,” *Enzyme Microb Technol*, vol. 26, p. 1316, 2000.

[20] R. G. Jensen, “Characteristics of the lipase from the mold, geotrichum candidum: a review,” *Lipids*, vol. 9, no. 3, pp. 149–57, 1974.

[21] P. Andersch, J. Berger, M. Hermann, K. Laumen, M. Lobell, R. Seemayer, and C. S. M. Waldinger, “Ester synthesis via acyl transfer (transesterification),” *Methods Enzymol*, vol. 286, pp. 406–443, 1997.

[22] U. Bornscheuer, “Strategies for improving the lipase catalysed preparation of chiral compounds,” in *Methods in non-aqueous enzymology*, M. Gupta, Ed. Basel: Birkhauser, 2000, pp. 90–109.

[23] S.-J. Kuo and K. Parkin, “Substrate preferences for lipase mediated acyl-exchange reactions with butteroil are concentration-dependent,” *J Am Oil Chem Soc*, vol. 70, pp. 393–9, 1993.

[24] C.-S. Chen and C. Sih, “General aspects and optimization of enantioselective biocatalysis in organic solvents: the use of lipases,” *Angew Chem Int Ed Engl*, vol. 28, pp. 695–708, 1989.

[25] A. Gutman and M. Shapira, “Synthetic applications of enzymatic reactions in organic solvents,” in *Advances in biochemical engineering and biotechnology*, F. A, Ed. Heidelberg: Springer, 1995, vol. 52, pp. 87–128.

[26] R. D. Joerger and M. J. Haas, “Alteration of chain length selectivity of a rhizopus delemar lipase through site-directed mutagenesis,” *Lipids*, vol. 29, no. 6, pp. 377–84, 1994.

[27] G. H. Peters, D. M. van Aalten, A. Svendsen, and R. Bywater, “Essential dynamics of lipase binding sites: the effect of inhibitors of different chain length,” *Protein Eng*, vol. 10, no. 2, pp. 149–58, 2005.

[28] J. Schmitt, S. Brocca, R. D. Schmid, and J. Pleiss, “Blocking the tunnel: engineering of candida rugosa lipase mutants with short chain length specificity,” *Protein Eng*, vol. 15, no. 7, pp. 595–601, 2002.

[29] J. Yang, Y. Koga, H. Nakano, and T. Yamane, “Modifying the chain-length selectivity of the lipase from burkholderia cepacia kwi-56 through in vitro combinatorial mutagenesis in the substrate-binding site,” *Protein Eng*, vol. 15, no. 2, pp. 147–52, 2002.

[30] J. D. Schrag and M. Cygler, “Lipases and alpha/beta hydrolase fold,” *Methods Enzymol*, vol. 284, pp. 85–107, 1997.

[31] M. Nardini and B. W. Dijkstra, “Alpha/beta hydrolase fold enzymes: the family keeps growing,” *Curr Opin Struct Biol*, vol. 9, no. 6, pp. 732–7, 1999.

[32] M. Holmquist, “Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms,” *Curr Protein Pept Sci*, vol. 1, no. 2, pp. 209–35, 2000.

[33] C. Carrasco-Lopez, C. Godoy, B. de Las Rivas, G. Fernandez-Lorente, J. M. Palomo, J. M. Guisan, R. Fernandez-Lafuente, M. Martinez-Ripoll, and J. A. Hermoso, “Activation of bacterial thermoalkalophilic lipases is spurred by dramatic structural rearrangements,” *J Biol Chem*, vol. 284, no. 7, pp. 4365–72, 2009.

[34] H. Wong and M. C. Schotz, “The lipase gene family,” *J Lipid Res*, vol. 43, no. 7, pp. 993–9, 2002.

[35] F. K. Winkler, A. D’Arcy, and W. Hunziker, “Structure of human pancreatic lipase,” *Nature*, vol. 343, no. 6260, pp. 771–4, 1990.

[36] G. G. Dodson, D. M. Lawson, and F. K. Winkler, “Structural and evolutionary relationships in lipase mechanism and activation,” *Faraday Discuss*, no. 93, pp. 95–105, 1992.

[37] A. M. Brzozowski, U. Derewenda, Z. S. Derewenda, G. G. Dodson, D. M. Lawson, J. P. Turkenburg, F. Bjorkling, B. Huge-Jensen, S. A. Patkar, and L. Thim, “A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex,” *Nature*, vol. 351, no. 6326, pp. 491–4, 1991.

[38] M. Cygler and J. D. Schrag, “Structure as basis for understanding interfacial properties of lipases,” *Methods Enzymol*, vol. 284, pp. 3–27, 1997.

[39] A. Hjorth, F. Carriere, C. Cudrey, H. Woldike, E. Boel, D. M. Lawson, F. Ferrato, C. Cambillau, G. G. Dodson, L. Thim, and et al., “A structural domain (the lid) found in pancreatic lipases is absent in the guinea pig (phospho)lipase,” *Biochemistry*, vol. 32, no. 18, pp. 4702–7, 1993.

[40] D. A. Lang, M. L. Mannesse, G. H. de Haas, H. M. Verheij, and B. W. Dijkstra, “Structural basis of the chiral selectivity of pseudomonas cepacia lipase,” *Eur J Biochem*, vol. 254, no. 2, pp. 333–40, 1998.

[41] J. Pleiss, M. Fischer, and R. D. Schmid, “Anatomy of lipase binding sites: the scissile fatty acid binding site,” *Chem Phys Lipids*, vol. 93, no. 1-2, pp. 67–80, 1998.

[42] R. Muralidhar, R. Chirumamilla, R. Marchant, V. Ramachandran, O. Ward, and P. Nigam, "Understanding lipase stereoselectivity," *WORLD JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY*, vol. 18, no. 2, pp. 81–97, 2002.

[43] L. Sarda and P. Desnuelle, "[actions of pancreatic lipase on esters in emulsions]," *Biochim Biophys Acta*, vol. 30, no. 3, pp. 513–21, 1958.

[44] U. Derewenda, A. M. Brzozowski, D. M. Lawson, and Z. S. Derewenda, "Catalysis at the interface: the anatomy of a conformational change in a triglyceride lipase," *Biochemistry*, vol. 31, no. 5, pp. 1532–41, 1992.

[45] R. Verger and G. Haas, "Interfacial enzyme kinetics of lipolysis," *Annual Review of Biophysics and Bioengineering*, vol. 5:, pp. 77–117, 1976.

[46] S. Labourdenne, A. Cagna, B. Delorme, G. Esposito, R. Verger, and C. Riviere, "Oil-drop tensiometer: applications for studying the kinetics of lipase action," *Methods Enzymol*, vol. 286, pp. 306–26, 1997.

[47] S. Ransac, M. Ivanova, I. Panaiotov, and R. Verger, "Monolayer techniques for studying lipase kinetics," *Methods Mol Biol*, vol. 109, pp. 279–302, 1999.

[48] P. Woolley, S. B. Petersen, and N. industrifond., *Lipases : their structure, biochemistry, and application*. Cambridge University Press, 1994.

[49] K. E. Jaeger and M. T. Reetz, "Microbial lipases form versatile tools for biotechnology," *Trends Biotechnol*, vol. 16, no. 9, pp. 396–403, 1998.

[50] F. X. Malcata and N. A. T. O. S. A. Division., *Engineering of/with lipases*, ser. NATO ASI series. Series E, Applied sciences. Dordrecht ; Boston: Kluwer Academic, 1996.

[51] A. Louwrier, “Industrial products - the return to carbohydrate-based industries,” *Biotechnology and Applied Biochemistry*, vol. 27, pp. 1–8, 1998.

[52] U. T. Bornscheuer, “Trends and challenges in enzyme technology,” *Adv Biochem Eng Biotechnol*, vol. 100, pp. 181–203, 2005.

[53] F. Bjorkling, S. E. Godtfredsen, and O. Kirk, “The future-impact of industrial lipases,” *Trends in Biotechnology*, vol. 9, no. 10, pp. 360–363, 1991.

[54] U. T. Bornscheuer, “Lipase-catalyzed syntheses of monoacylglycerols,” *Enzyme and Microbial Technology*, vol. 17, no. 7, pp. 578–586, 1995.

[55] B. Borgstrm and H. L. Brockman, *Lipases*. Amsterdam ; New York New York, N.Y.: Elsevier ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co., 1984.

[56] M. L. Rua, C. Schmidt-Dannert, S. Wahl, A. Sprauer, and R. D. Schmid, “Thermoalkalophilic lipase of bacillus thermocatenulatus large-scale production, purification and properties: aggregation behaviour and its effect on activity,” *J Biotechnol*, vol. 56, no. 2, pp. 89–102, 1997.

[57] E. W. Seitz, “Industrial application of microbial lipases: a review,” *Journal of the American Oil Chemists Society*, vol. 51, no. 2, pp. 12–6, 1974.

[58] M. Schoemaker, R. Feldbrugge, B. Grundig, and F. Spener, “The lipoxygenase sensor, a new approach in essential fatty acid determination in foods,” *Biosens Bioelectron*, vol. 12, no. 11, pp. 1089–99, 1997.

[59] K. E. Jaeger and T. Eggert, “Lipases for biotechnology,” *Curr Opin Biotechnol*, vol. 13, no. 4, pp. 390–7, 2002.

[60] M. T. Reetz, "Lipases as practical biocatalysts," *Curr Opin Chem Biol*, vol. 6, no. 2, pp. 145–50, 2002.

[61] L. Poppe and L. Novk, *Selective biocatalysis : a synthetic approach*. Weinheim ; New York: VCH, 1992.

[62] K. Drauz and H. Waldmann, *Enzyme catalysis in organic synthesis : a comprehensive handbook*, 2nd ed. Weinheim ; New York: Wiley-VCH, 2002.

[63] M. G. van Oort, A. M. Deveer, R. Dijkman, M. L. Tjeenk, H. M. Verheij, G. H. de Haas, E. Wenzig, and F. Gotz, "Purification and substrate specificity of staphylococcus hyicus lipase," *Biochemistry*, vol. 28, no. 24, pp. 9278–85, 1989.

[64] E. J. Gilbert, "Pseudomonas lipases: biochemical properties and molecular cloning," *Enzyme Microb Technol*, vol. 15, no. 8, pp. 634–45, 1993.

[65] P. Grochulski, Y. Li, J. D. Schrag, F. Bouthillier, P. Smith, D. Harrison, B. Rubin, and M. Cygler, "Insights into interfacial activation from an open structure of candida rugosa lipase," *J Biol Chem*, vol. 268, no. 17, pp. 12 843–7, 1993.

[66] K. E. Jaeger, S. Ransac, B. W. Dijkstra, C. Colson, M. van Heuvel, and O. Misset, "Bacterial lipases," *FEMS Microbiol Rev*, vol. 15, no. 1, pp. 29–63, 1994.

[67] T. Godfrey and S. West, *Industrial enzymology*, 2nd ed. New York: Stockton Press, 1996.

[68] M. Boston, C. Requadt, S. Danko, A. Jarnagin, E. Ashizawa, S. Wu, A. J. Poulose, and R. Bott, "Structure and function engineered pseudomonas mendocina lipase," *Methods Enzymol*, vol. 284, pp. 298–317, 1997.

[69] B. Rubin and E. Dennis, "Methods in enzymology: Lipases, part a: Biotechnology," 1997.

[70] H. Wong, R. C. Davis, J. S. Hill, D. Yang, and M. C. Schotz, “Lipase engineering: a window into structure-function relationships,” *Methods Enzymol*, vol. 284, pp. 171–84, 1997.

[71] R. Gupta, N. Gupta, and P. Rathi, “Bacterial lipases: an overview of production, purification and biochemical properties,” *Appl Microbiol Biotechnol*, vol. 64, no. 6, pp. 763–81, 2004.

[72] J. J. Tiesinga, G. van Pouderoyen, M. Nardini, S. Ransac, and B. W. Dijkstra, “Structural basis of phospholipase activity of staphylococcus hyicus lipase,” *J Mol Biol*, vol. 371, no. 2, pp. 447–56, 2007.

[73] J. S. White and D. C. White, *Source book of enzymes*. Boca Raton, Fla.: CRC Press, 1997.

[74] A. Tomasini, G. Bustillo, and J. Lebeault, “Fat lipolyzed with a commercial lipase for the production of blue cheese flavour,” *International Dairy Journal*, vol. 3, no. 2, pp. 117–127, 1993.

[75] R. Farrell, K. Hata, and M. . Wall, “Solving pitch problems in pulp and paper processes by the use of enzymes or fungi,” *Advances in Biochemical Engineering/Biotechnology*, vol. 57, p. 197212, 1997.

[76] P. Bajpai, “Application of enzymes in the pulp and paper industry,” *Biotechnol Prog*, vol. 15, no. 2, pp. 147–57, 1999.

[77] A. Gutierrez, J. C. del Rio, M. J. Martinez, and A. T. Martinez, “The biotechnological control of pitch in paper pulp manufacturing,” *Trends Biotechnol*, vol. 19, no. 9, pp. 340–8, 2001.

[78] R. Margesin and F. Schinner, *Biotechnological applications of cold-adapted organisms*. Berlin ; New York: Springer, 1999.

[79] R. N. Patel, “Microbial/enzymatic synthesis of chiral drug intermediates,” *Adv Appl Microbiol*, vol. 47, pp. 33–78, 2000.

[80] P. Rasor, E. Voss, and P. H. Baartmans, “Enzyme-catalyzed processes in pharmaceutical industry,” *Applied Catalysis A: General*, vol. 221, pp. 145–158, 2001.

[81] B. Joseph, P. W. Ramteke, and G. Thomas, “Cold active microbial lipases: some hot issues and recent developments,” *Biotechnol Adv*, vol. 26, no. 5, pp. 457–70, 2008.

[82] R. Piamtongkam, S. Duquesne, F. Bordes, S. Barbe, I. Andre, A. Marty, and W. Chulalaksananukul, “Enantioselectivity of candida rugosa lipases (lip1, lip3, and lip4) towards 2-bromo phenylacetic acid octyl esters controlled by a single amino acid,” *Biotechnol Bioeng*, vol. 108, no. 8, pp. 1749–56, 2011.

[83] F. Hasan, A. Shah, and A. Hameed, “Industrial applications of microbial lipases,” *Enzyme Microb Technol*, vol. 39, p. 23551, 2006.

[84] G. M. Kempler, “Production of flavor compounds by microorganisms,” *Adv Appl Microbiol*, vol. 29, pp. 29–51, 1983.

[85] S. Bloomer, P. Adlercreutz, and B. Mattiasson, “Triglyceride interesterification by lipases. 1. cocoa butter equivalents from a fraction of palm oil,” *J Am Oil Chem Soc*, vol. 67, pp. 519–24, 1990.

[86] S. Krishna and N. Karanth, “Lipases and lipase-catalyzed esterification reactions in nonaqueous media,” *Catal Rev*, vol. 44, pp. 499–591, 2002.

[87] J. Chopineau, F. D. McCafferty, M. Therisod, and A. M. Klibanov, “Production of biosurfactants from sugar alcohols and vegetable oils catalyzed by lipases in a nonaqueous medium,” *Biotechnol Bioeng*, vol. 31, no. 3, pp. 208–14, 1988.

[88] M. Iso, B. Chen, M. Eguchi, T. Kudo, and S. Shrestha, "Production of biodiesel fuel from triglycerides and alcohol using immobilized lipase," *Journal of Molecular Catalysis B: Enzymatic*, vol. 16, no. 1, pp. 53–58, 2001.

[89] P. S. Bisen, B. S. Sanodiya, G. S. Thakur, R. K. Baghel, and G. B. Prasad, "Biodiesel production with special emphasis on lipase-catalyzed transesterification," *Biotechnol Lett*, vol. 32, no. 8, pp. 1019–30, bisen, Prakash S year = 2010.

[90] D. Lee, Y. Koh, K. Kim, B. Kim, H. Choi, D. Kim, M. T. Suhartono, and Y. Pyun, "Isolation and characterization of a thermophilic lipase from bacillus thermoleovorans id-1," *FEMS Microbiol Lett*, vol. 179, no. 2, pp. 393–400, 1999.

[91] A. L. Margolin, "Enzymes in the synthesis of chiral drugs," *Enzyme Microb Technol*, vol. 15, no. 4, pp. 266–80, 1993.

[92] U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, and K. Robins, "Engineering the third wave of biocatalysis," *Nature*, vol. 485, no. 7397, pp. 185–94, 2012.

[93] F. T. Robb, *Thermophiles : biology and technology at high temperatures*. Boca Raton, FL: CRC Press, 2008.

[94] P. R. Carey, *Protein engineering and design*. San Diego, Calif.: Academic Press, 1996.

[95] J. D. Hirst, M. Vieth, J. Skolnick, and r. Brooks, C. L., "Predicting leucine zipper structures from sequence," *Protein Eng*, vol. 9, no. 8, pp. 657–62, 1996.

[96] P. V. Nikolova, J. Henckel, D. P. Lane, and A. R. Fersht, "Semirational design of active tumor suppressor p53 dna binding domain with enhanced stability," *Proc Natl Acad Sci U S A*, vol. 95, no. 25, pp. 14 675–80, 1998.

[97] T. Kawabata, M. Ota, and K. Nishikawa, “The protein mutant database.”

[98] F. H. Arnold, “Engineering proteins for nonnatural environments,” *FASEB J*, vol. 7, no. 9, pp. 744–9, 1993.

[99] K. Chen and F. H. Arnold, “Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin e for catalysis in dimethylformamide,” *Proc Natl Acad Sci U S A*, vol. 90, no. 12, pp. 5618–22, 1993.

[100] L. E. Nagy, J. E. Meuser, S. Plummer, M. Seibert, M. L. Ghirardi, P. W. King, D. Ahmann, and M. C. Posewitz, “Application of gene-shuffling for the rapid generation of novel [fefe]-hydrogenase libraries,” *Biotechnol Lett*, vol. 29, no. 3, pp. 421–30, 2007.

[101] Y. Yu and S. Lutz, “Circular permutation: a different way to engineer enzyme structure and function,” *Trends Biotechnol*, vol. 29, no. 1, pp. 18–25, 2011.

[102] S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease, “Site-directed mutagenesis by overlap extension using the polymerase chain reaction,” *Gene*, vol. 77, no. 1, pp. 51–9, 1989.

[103] S. K. Padhi, D. J. Bougioukou, and J. D. Stewart, “Site-saturation mutagenesis of tryptophan 116 of saccharomyces pastorianus old yellow enzyme uncovers stereocomplementary variants,” *J Am Chem Soc*, vol. 131, no. 9, pp. 3271–80, 2009.

[104] R. Guerois and M. Lpez de la Paz, *Protein design : methods and applications*, ser. Methods in molecular biology,. Totowa, N.J.: Humana Press, 2006.

[105] F. C. Bernstein, T. F. Koetzle, G. J. Williams, J. Meyer, E. F., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, “The protein data bank. a computer-based archival file for macromolecular structures,” *Eur J Biochem*, vol. 80, no. 2, pp. 319–24, 1977.

[106] I. Y. Torshin, *Bioinformatics in the post-genomic era : the role of biophysics*. New York: Nova Biomedical Books, 2006.

[107] A. A. Lazakidou, *Biocomputation and biomedical informatics : case studies and applications*. Medical Information Science Reference, 2010.

[108] S. J. Park and J. R. Cochran, *Protein engineering and design*, ser. London mathematical society, student texts. Boca Raton: CRC Press, 2010.

[109] C. Jarzynski, "Nonequilibrium equality for free energy differences," *Phys Rev Lett*, vol. 78, p. 26902693, 1997.

[110] J. Kua, Y. Zhang, and J. A. McCammon, "Studying enzyme binding specificity in acetylcholinesterase using a combined molecular dynamics and multiple docking approach," *J Am Chem Soc*, vol. 124, no. 28, pp. 8260–7, 2002.

[111] T. Frembgen-Kesner and A. H. Elcock, "Computational sampling of a cryptic drug binding site in a protein receptor: explicit solvent molecular dynamics and inhibitor docking to p38 map kinase," *J Mol Biol*, vol. 359, no. 1, pp. 202–14, 2006.

[112] G. Morris, D. Goodsell, R. Halliday, R. Huey, W. Hart, R. Belew, and A. Olson, "Automated docking using a lamarckian genetic algorithm and and empirical binding free energy function," *J Comput Chem*, no. 19, 1998.

[113] I. Schellhammer and M. Rarey, "Flexx-scan: fast, structure-based virtual screening," *Proteins*, vol. 57, no. 3, pp. 504–17, 2004.

[114] S. J. de Vries, M. van Dijk, and A. M. Bonvin, "The haddock web server for data-driven biomolecular docking," *Nat Protoc*, vol. 5, no. 5, pp. 883–97, 2010.

[115] O. Trott and A. J. Olson, "Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *Journal of Computational Chemistry*, vol. 31, no. 2, pp. 455–61, 2010.

[116] S. F. Sousa, P. A. Fernandes, and M. J. Ramos, "Protein-ligand docking: current status and future challenges," *Proteins*, vol. 65, no. 1, pp. 15–26, 2006.

[117] W. Wang and P. A. Kollman, "Computational study of protein specificity: the molecular basis of hiv-1 protease drug resistance," *Proc Natl Acad Sci U S A*, vol. 98, no. 26, pp. 14 937–42, 2001.

[118] P. Ju, G. Pages, R. P. Riek, P. C. Chen, A. M. Torres, P. S. Bansal, S. Kuyucak, P. W. Kuchel, and J. I. Vandenberg, "The pore domain outer helix contributes to both activation and inactivation of the herg k+ channel," *J Biol Chem*, vol. 284, no. 2, pp. 1000–8, 2009.

[119] T. Werner, K. Sander, Y. Tanrikulu, T. Kottke, E. Proschak, H. Stark, and G. Schneider, "In silico characterization of ligand binding modes in the human histamine h4 receptor and their impact on receptor activation," *Chembiochem*, vol. 11, no. 13, pp. 1850–5, 1997.

[120] J. A. McCammon, B. R. Gelin, and M. Karplus, "Dynamics of folded proteins," *Nature*, vol. 267, no. 5612, pp. 585–90, 1977.

[121] P. K. Warme and H. A. Scheraga, "Refinement of the x-ray structure of lysozyme by complete energy minimization," *Biochemistry*, vol. 13, no. 4, pp. 757–67, 1974.

[122] A. T. Hagler and S. Lifson, "Energy functions for peptides and proteins. ii. the amide hydrogen bond and calculation of amide crystal properties," *J Am Chem Soc*, vol. 96, no. 17, pp. 5327–35, 1974.

[123] A. T. Hagler, E. Huler, and S. Lifson, “Energy functions for peptides and proteins. i. derivation of a consistent force field including the hydrogen bond from amide crystals,” *J Am Chem Soc*, vol. 96, no. 17, pp. 5319–27, 1974.

[124] J. Hermans and J. McQueen, “Computer manipulation of (macro)molecules with the method of local change,” *Acta Cryst.*, vol. A30, pp. 730–739, 1974.

[125] N. Allinger, “Conformational analysis. 130. mm2. a hydrocarbon force field utilizing v1 and v2 torsional terms,” *J. Am. Chem. Soc.*, vol. 99, no. 25, pp. 8127–8134, 1977.

[126] B. R. Brooks, r. Brooks, C. L., J. Mackerell, A. D., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, “Charmm: the biomolecular simulation program,” *Journal of Computational Chemistry*, vol. 30, no. 10, pp. 1545–614, 2009.

[127] S. Weiner, P. Kollman, D. Nguyen, and D. Case, “An all atom force field for simulations of proteins and nucleic acids,” *J. Comp. Chem*, vol. 7, pp. 230–252, 1986.

[128] C. D. Christ, A. E. Mark, and W. F. van Gunsteren, “Basic ingredients of free energy calculations: a review,” *J Comput Chem*, vol. 31, no. 8, pp. 1569–82, 2010.

[129] M. Karplus and J. A. McCammon, “Molecular dynamics simulations of biomolecules,” *Nat Struct Biol*, vol. 9, no. 9, pp. 646–52, 2002.

[130] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten, “Scalable molecular dynamics with namd,” *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1781–802, 2005.

[131] C. Chipot and A. Pohorille, *Free energy calculations : theory and applications in chemistry and biology*, study ed., ser. Springer series in chemical physics. New York: Springer, 2007.

[132] S. K. Ludemann, V. Lounnas, and R. C. Wade, "How do substrates enter and products exit the buried active site of cytochrome p450cam? 2. steered molecular dynamics and adiabatic mapping of substrate pathways," *J Mol Biol*, vol. 303, no. 5, pp. 813–30, 2000.

[133] Y. Gu, I. H. Shrivastava, S. G. Amara, and I. Bahar, "Molecular simulations elucidate the substrate translocation pathway in a glutamate transporter," *Proc Natl Acad Sci U S A*, vol. 106, no. 8, pp. 2589–94, 2009.

[134] V. Guallar, C. Lu, K. Borrelli, T. Egawa, and S. R. Yeh, "Ligand migration in the truncated hemoglobin-ii from mycobacterium tuberculosis: the role of g8 tryptophan," *J Biol Chem*, vol. 284, no. 5, pp. 3106–16, 2009.

[135] K. J. McLean, P. Lafite, C. Levy, M. R. Cheesman, N. Mast, I. A. Pikuleva, D. Leys, and A. W. Munro, "The structure of mycobacterium tuberculosis cyp125: molecular basis for cholesterol binding in a p450 needed for host infection," *J Biol Chem*, vol. 284, no. 51, pp. 35 524–33, 2009.

[136] M. Perakyla, "Ligand unbinding pathways from the vitamin d receptor studied by molecular dynamics simulations," *Eur Biophys J*, vol. 38, no. 2, pp. 185–98, 2009.

[137] P. Pongprayoon, O. Beckstein, C. L. Wee, and M. S. Sansom, "Simulations of anion transport through oprp reveal the molecular basis for high affinity and selectivity for phosphate," *Proc Natl Acad Sci U S A*, vol. 106, no. 51, pp. 21 614–8, 2009.

[138] D. P. Claxton, M. Quick, L. Shi, F. D. de Carvalho, H. Weinstein, J. A. Javitch, and H. S. McHaourab, "Ion/substrate-dependent conformational dynamics of a bacte-

rial homolog of neurotransmitter:sodium symporters," *Nat Struct Mol Biol*, vol. 17, no. 7, pp. 822–9, 2010.

[139] B. Knapp, U. Omasits, W. Schreiner, and M. M. Epstein, "A comparative approach linking molecular dynamics of altered peptide ligands and mhc with in vivo immune responses," *PLoS One*, vol. 5, no. 7, p. e11653, 2010.

[140] L. Bu, G. T. Beckham, M. R. Shirts, M. R. Nimlos, W. S. Adney, M. E. Himmel, and M. F. Crowley, "Probing carbohydrate product expulsion from a processive cellulase with multiple absolute binding free energy methods," *J Biol Chem*, vol. 286, no. 20, pp. 18 161–9, 2011.

[141] A. Pesce, M. Nardini, S. Dewilde, L. Capece, M. A. Marti, S. Congia, M. D. Salter, G. C. Blouin, D. A. Estrin, P. Ascenzi, L. Moens, M. Bolognesi, and J. S. Olson, "Ligand migration in the apolar tunnel of cerebratulus lacteus mini-hemoglobin," *J Biol Chem*, vol. 286, no. 7, pp. 5347–58, 2011.

[142] B. Isralewitz, J. Baudry, J. Gullingsrud, D. Kosztin, and K. Schulten, "Steered molecular dynamics investigations of protein function," *J Mol Graph Model*, vol. 19, no. 1, pp. 13–25, 2001.

[143] S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, "Free energy calculation from steered molecular dynamics simulations using jarzynskis equality," *J Chem Phys*, vol. 119, pp. 3559–3566, 2003.

[144] M. S. Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamante, "Folding-unfolding transitions in single titin molecules characterized with laser tweezers," *Science*, vol. 276, no. 5315, pp. 1112–6, 1997.

[145] J. B. Thompson, H. G. Hansma, P. K. Hansma, and K. W. Plaxco, “The backbone conformational entropy of protein folding: experimental measures from atomic force microscopy,” *J Mol Biol*, vol. 322, no. 3, pp. 645–52, 2002.

[146] D. A. Pearlman, “A comparison of alternative approaches to free-energy calculations,” *Journal of Physical Chemistry*, vol. 98, no. 5, pp. 1487–1493, 1994.

[147] M. T. Sonoda, L. Martinez, P. Webb, M. S. Skaf, and I. Polikarpov, “Ligand dissociation from estrogen receptor is mediated by receptor dimerization: evidence from molecular dynamics simulations,” *Mol Endocrinol*, vol. 22, no. 7, pp. 1565–78, 2008.

[148] T. Bastug and S. Kuyucak, “Application of jarzynskis equality in simple versus complex systems,” *Chemical Physics Letters*, vol. 436, p. 38338, 2007.

[149] T. Bastug, P. C. Chen, S. M. Patra, and S. Kuyucak, “Potential of mean force calculations of ligand binding to ion channels from jarzynski’s equality and umbrella sampling,” *J Chem Phys*, vol. 128, no. 15, p. 155104, 2008.

[150] T. Bastug and S. Kuyucak, “Energetics of ion permeation, rejection, binding, and block in gramicidin a from free energy simulations,” *Biophys J*, vol. 90, no. 11, pp. 3941–50, 2006.

[151] ——, “Free energy simulations of single and double ion occupancy in gramicidin a,” *J Chem Phys*, vol. 126, no. 10, p. 105103, 2007.

[152] G. E. Crooks, “Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems,” *Journal of Statistical Physics*, vol. 90, no. 5-6, pp. 1481–1487, 1998.

[153] H. H. Hogrefe, J. Cline, G. L. Youngblood, and R. M. Allen, “Creating randomized amino acid libraries with the quikchange multi site-directed mutagenesis kit,” *Biotechniques*, vol. 33, no. 5, pp. 1158–60, 1162, 1164–5, 2002.

[154] M. Chihara-Siomi, K. Yoshikawa, N. Oshima-Hirayama, K. Yamamoto, Y. Sogabe, T. Nakatani, T. Nishioka, and J. Oda, “Purification, molecular cloning, and expression of lipase from pseudomonas aeruginosa,” *Arch Biochem Biophys*, vol. 296, no. 2, pp. 505–13, 1992.

[155] C. Schmidt-Dannert, M. L. Rua, and R. D. Schmid, “Two novel lipases from thermophile bacillus thermocatenulatus: screening, purification, cloning, overexpression, and properties,” *Methods Enzymol*, vol. 284, pp. 194–220, 1997.

[156] S. T. Jeong, H. K. Kim, S. J. Kim, S. W. Chi, J. G. Pan, T. K. Oh, and S. E. Ryu, “Novel zinc-binding center and a temperature switch in the bacillus stearothermophilus l1 lipase,” *J Biol Chem*, vol. 277, no. 19, pp. 17 041–7, 2002.

[157] H. Kirino, M. Aoki, M. Aoshima, Y. Hayashi, M. Ohba, A. Yamagishi, T. Wakagi, and T. Oshima, “Hydrophobic interaction at the subunit interface contributes to the thermostability of 3-isopropylmalate dehydrogenase from an extreme thermophile, thermus thermophilus,” *European Journal of Biochemistry*, vol. 220, no. 1, pp. 275–281, 1994.

[158] L. F. Delboni, S. C. Mande, F. Rentier-Delrue, V. Mainfroid, S. Turley, F. M. Vellieux, J. A. Martial, and W. G. Hol, “Crystal structure of recombinant triosephosphate isomerase from bacillus stearothermophilus. an analysis of potential thermostability factors in six isomerases with known three-dimensional structures points to the importance of hydrophobic interactions,” *Protein Sci*, vol. 4, no. 12, pp. 2594–604, 1995.

[159] V. Z. Spassov, A. D. Karshikoff, and R. Ladenstein, “The optimization of protein-solvent interactions: thermostability and the role of hydrophobic and electrostatic interactions,” *Protein Sci*, vol. 4, no. 8, pp. 1516–27, 1995.

[160] J. L. Arpigny and K. E. Jaeger, “Bacterial lipolytic enzymes: classification and properties,” *Biochem J*, vol. 343 Pt 1, pp. 177–83, 1999.

[161] H. K. Kim, S. Y. Park, J. K. Lee, and T. K. Oh, “Gene cloning and characterization of thermostable lipase from bacillus stearothermophilus l1,” *Biosci Biotechnol Biochem*, vol. 62, no. 1, pp. 66–71, 1998.

[162] D. T. Quyen, C. Schmidt-Dannert, and R. D. Schmid, “High-level expression of a lipase from bacillus thermocatenulatus btl2 in pichia pastoris and some properties of the recombinant lipase,” *Protein Expr Purif*, vol. 28, no. 1, pp. 102–10, 2003.

[163] G. Fernandez-Lorente, C. A. Godoy, A. A. Mendes, F. Lopez-Gallego, V. Grazu, B. de Las Rivas, J. M. Palomo, J. Hermoso, R. Fernandez-Lafuente, and J. M. Guisan, “Solid-phase chemical amination of a lipase from bacillus thermocatenulatus to improve its stabilization via covalent immobilization on highly activated glyoxyl-agarose,” *Biomacromolecules*, vol. 9, no. 9, pp. 2553–61, 2008.

[164] M. Tayyab, N. Rashid, and M. Akhtar, “Isolation and identification of lipase producing thermophilic geobacillus sp. sbs-4s: cloning and characterization of the lipase,” *J Biosci Bioeng*, vol. 111, no. 3, pp. 272–8, 2011.

[165] A. Schrodel, J. Volz, and A. de Marco, “Fusion tags and chaperone co-expression modulate both the solubility and the inclusion body features of the recombinant clipb14 serine protease,” *J Biotechnol*, vol. 120, no. 1, pp. 2–10, 2005.

[166] S. Hirose, Y. Kawamura, M. Mori, K. Yokota, T. Noguchi, and N. Goshima, "Development and evaluation of data-driven designed tags (ddts) for controlling protein solubility," *N Biotechnol*, vol. 28, no. 3, pp. 225–31, 2011.

[167] M. M. Islam, M. A. Khan, and Y. Kuroda, "Analysis of amino acid contributions to protein solubility using short peptide tags fused to a simplified bpti variant," *Biochim Biophys Acta*, vol. 1824, no. 10, pp. 1144–50, 2012.

[168] E. S. Hemdan, Y. J. Zhao, E. Sulkowski, and J. Porath, "Surface topography of histidine residues: a facile probe by immobilized metal ion affinity chromatography," *Proc Natl Acad Sci U S A*, vol. 86, no. 6, pp. 1811–5, 1989.

[169] J. E. Tropea, S. Cherry, and D. S. Waugh, "Expression and purification of soluble his(6)-tagged tev protease," *Methods Mol Biol*, vol. 498, pp. 297–307, 2009.

[170] C. Schmidt-Dannert, H. Sztajer, W. Stocklein, U. Menge, and R. D. Schmid, "Screening, purification and properties of a thermophilic lipase from bacillus thermocatenulatus," *Biochim Biophys Acta*, vol. 1214, no. 1, pp. 43–53, 1994.

[171] J. Newman, "Novel buffer systems for macromolecular crystallization," *Acta Crystallogr D Biol Crystallogr*, vol. 60, no. Pt 3, pp. 610–2, 2004.

[172] D. L. Ollis, E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, and et al., "The alpha/beta hydrolase fold," *Protein Eng*, vol. 5, no. 3, pp. 197–211, 1992.

[173] B. Zentgraf and T. J. Ahern, "Practical importance of enzyme stability," *Pure Appl. Chem.*, vol. 63, no. 10, pp. 1527–1540, 1991.

[174] K. Hult and M. Holmquist, "Kinetic, molecular modeling and synthetic applications with microbial lipases," *Methods Enzymol*, vol. 286, pp. 246–60, 1997.

[175] G. D. Haki and S. K. Rakshit, “Developments in industrially important thermostable enzymes: a review,” *Bioresour Technol*, vol. 89, no. 1, pp. 17–34, 2003.

[176] D. W. Armstrong and H. Yamazaki, “Natural flavours production: a biotechnological approach,” *Trends in Biotechnology*, vol. 4, no. 10, p. 264267, 1986.

[177] Z. X. Chen and T. R. Breitman, “Tributyrin: a prodrug of butyric acid for potential clinical application in differentiation therapy,” *Cancer Res*, vol. 54, no. 13, pp. 3494–9, 1994.

[178] P. Durre, “Biobutanol: an attractive biofuel,” *Biotechnol J*, vol. 2, no. 12, pp. 1525–34, 2007.

[179] A. Rephaeli, S. Waks-Yona, A. Nudelman, I. Tarasenko, N. Tarasenko, D. R. Phillips, S. M. Cutts, and G. Kessler-Icekson, “Anticancer prodrugs of butyric acid and formaldehyde protect against doxorubicin-induced cardiotoxicity,” *Br J Cancer*, vol. 96, no. 11, pp. 1667–74, 2001.

[180] F. D. Gunstone, *Oils and fats in the food industry*, ser. Food industry briefing series. Oxford ; Ames, Iowa: Wiley-Blackwell Pub., 2008.

[181] C.-T. Ho and R. S. of Chemistry (Great Britain), *Recent advances in food and flavor chemistry : food flavors and encapsulation, health benefits, analytical methods, and molecular biology of functional foods*, ser. Special publication. Cambridge: RSC Pub., 2010.

[182] C. Shu, J. Cai, L. Huang, X. Zhu, and Z. Xu, “Biocatalytic production of ethyl butyrate from butyric acid with immobilized candida rugosa lipase on cotton cloth,” *Journal of Molecular Catalysis B: Enzymatic*, vol. 72, no. 3-4, pp. 139–144, 2011.

[183] J. M. Palomo, R. L. Segura, G. Fernandez-Lorente, M. Pernas, M. L. Rua, J. M. Guisan, and R. Fernandez-Lafuente, “Purification, immobilization, and stabilization

of a lipase from bacillus thermocatenulatus by interfacial adsorption on hydrophobic supports," *Biotechnol Prog*, vol. 20, no. 2, pp. 630–5, 2004.

[184] H. Matsumura, T. Yamamoto, T. C. Leow, T. Mori, A. B. Salleh, M. Basri, T. Inoue, Y. Kai, and R. N. Rahman, "Novel cation-pi interaction revealed by crystal structure of thermoalkalophilic lipase," *Proteins*, vol. 70, no. 2, pp. 592–8, 2008.

[185] T. Maniatis, E. F. Fritsch, and J. Sambrook, *Molecular cloning : a laboratory manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory, 1982.

[186] M. Dagert and S. D. Ehrlich, "Prolonged incubation in calcium chloride improves the competence of escherichia coli cells," *Gene*, vol. 6, no. 1, pp. 23–8, 1979.

[187] S. R. Maloy, V. J. Stewart, and R. K. Taylor, *Genetic analysis of pathogenic bacteria : a laboratory manual*. Plainview, N.Y.: Cold Spring Harbor Laboratory Press, 1996.

[188] G. Kouker and K. E. Jaeger, "Specific and sensitive plate assay for bacterial lipases," *Appl Environ Microbiol*, vol. 53, no. 1, pp. 211–3, 1987.

[189] M. M. Bradford, "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding," *Anal Biochem*, vol. 72, pp. 248–54, 1976.

[190] R. E. Wrolstad, *Current protocols in food analytical chemistry*. New York ;: Wiley, 2001.

[191] A. Fersht, *Enzyme structure and mechanism*, 2nd ed. New York: W.H. Freeman, 1985.

[192] A. R. Fersht, J. P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Waye, and G. Winter, "Hydrogen bonding and biological

specificity analysed by protein engineering," *Nature*, vol. 314, no. 6008, pp. 235–8, 1985.

[193] L. Graf, A. Jancso, L. Szilagyi, G. Hegyi, K. Pinter, G. Naray-Szabo, J. Hepp, K. Medzihradszky, and W. J. Rutter, "Electrostatic complementarity within the substrate-binding pocket of trypsin," *Proc Natl Acad Sci U S A*, vol. 85, no. 14, pp. 4961–5, 1988.

[194] P. C. Chen and S. Kuyucak, "Accurate determination of the binding free energy for kcsa-charybdotoxin complex from the potential of mean force calculations with restraints," *Biophys J*, vol. 100, no. 10, pp. 2466–74, 2011.

[195] L. Stols, M. Gu, L. Dieckman, R. Raffen, F. R. Collart, and M. I. Donnelly, "A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site," *Protein Expr Purif*, vol. 25, no. 1, pp. 8–15, 2002.

[196] M. Levisson, J. van der Oost, and S. W. Kengen, "Carboxylic ester hydrolases from hyperthermophiles," *Extremophiles*, vol. 13, no. 4, pp. 567–81, 2009.

[197] C. Schmidt-Dannert, M. L. Rua, H. Atomi, and R. D. Schmid, "Thermoalkalophilic lipase of bacillus thermocatenulatus. i. molecular cloning, nucleotide sequence, purification and some properties," *Biochim Biophys Acta*, vol. 1301, no. 1-2, pp. 105–14, 1996.

[198] L. Michaelis, M. L. Menten, K. A. Johnson, and R. S. Goody, "The original michaelis constant: translation of the 1913 michaelis-menten paper," *Biochemistry*, vol. 50, no. 39, pp. 8264–9, 2011.

[199] G. E. Briggs and J. B. Haldane, "A note on the kinetics of enzyme action," *Biochem J*, vol. 19, no. 2, pp. 338–9, 1925.

[200] M. Holmquist, M. Martinelle, I. G. Clausen, S. Patkar, A. Svendsen, and K. Hult, "Trp89 in the lid of humicola lanuginosa lipase is important for efficient hydrolysis of tributyrin," *Lipids*, vol. 29, no. 9, pp. 599–603, 1994.

[201] M. L. Rua, C. SchmidtDannert, S. Wahl, A. Sprauer, and R. D. Schmid, "Thermoalkalophilic lipase of bacillus thermocatenulatus large-scale production, purification and properties: aggregation behaviour and its effect on activity," *Journal of Biotechnology*, vol. 56, no. 2, pp. 89–102, 1997.

[202] J. Hayakawa, I. Kajihara, K. Morikawa, M. Oda, and Y. Fujio, "Denaturation of bovine serum albumin (bsa) and ovalbumin by high pressure, heat and chemicals," *Journal of Food Science*, vol. 57, no. 2, 1992.

[203] N. H. Schlieben, K. Niefind, and D. Schomburg, "Expression, purification, and aggregation studies of his-tagged thermoalkalophilic lipase from bacillus thermocatenulatus," *Protein Expr Purif*, vol. 34, no. 1, pp. 103–10, 2004.

[204] B. R. Brooks, r. Brooks, C. L., J. Mackerell, A. D., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, "Charmm: the biomolecular simulation program," *J Comput Chem*, vol. 30, no. 10, pp. 1545–614, 2009.

[205] J. A. D. MacKerell, D. Bashford, M. Bellott, R. Dunbrack Jr., J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, III, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, "All-atom empirical potential for

molecular modeling and dynamics studies of proteins," *Journal of Physical Chemistry B*, vol. 102, pp. 3586–3616, 1998.

[206] M. Feig, A. MacKerell, Jr., and C. Brooks, III, "Force field influence on the observation of a-helical protein structures in molecular dynamics simulations," *Journal of Physical Chemistry B*, vol. 107, pp. 2831–2836, 2003.

[207] A. MacKerell, Jr., M. Feig, and C. Brooks, "Improved treatment of the protein backbone in empirical force fields," *Journal of the American Chemical Society*, vol. 126, pp. 698–699, 2004.

[208] W. R. Kirk, E. Kurian, and F. G. Prendergast, "Characterization of the sources of protein-ligand affinity: 1-sulfonato-8-(1')anilinonaphthalene binding to intestinal fatty acid binding protein," *Biophys J*, vol. 70, no. 1, pp. 69–83, 1996.

[209] A. V. Pastukhov and I. J. Ropson, "Fluorescent dyes as probes to study lipid-binding proteins," *Proteins*, vol. 53, no. 3, pp. 607–15, 2003.

[210] S. Jana, T. K. Chaudhuri, and J. K. Deb, "Effects of guanidine hydrochloride on the conformation and enzyme activity of streptomycin adenylyltransferase monitored by circular dichroism and fluorescence spectroscopy," *Biochemistry (Mosc)*, vol. 71, no. 11, pp. 1230–7, 2006.

[211] Y. Cordeiro, L. M. Lima, M. P. Gomes, D. Foguel, and J. L. Silva, "Modulation of prion protein oligomerization, aggregation, and beta-sheet conversion by 4,4'-dianilino-1,1'-binaphthyl-5,5'-sulfonate (bis-ans)," *J Biol Chem*, vol. 279, no. 7, pp. 5346–52, 2004.

[212] S. Pedersen, L. Nesgaard, R. P. Baptista, E. P. Melo, S. R. Kristensen, and D. E. Otzen, "ph-dependent aggregation of cutinase is efficiently suppressed by 1,8-ans," *Biopolymers*, vol. 83, no. 6, pp. 619–29, 2006.

[213] V. Kayser, N. Chennamsetty, V. Voynov, B. Helk, and B. L. Trout, “Conformational stability and aggregation of therapeutic monoclonal antibodies studied with ans and thioflavin t binding,” *MAbs*, vol. 3, no. 4, pp. 408–11, 2011.

[214] G. D. Fasman, *Circular dichroism and the conformational analysis of biomolecules*, ser. The language of science. New York: Plenum Press, 1996.

[215] W. Kadima, A. McPherson, M. F. Dunn, and F. A. Jurnak, “Characterization of precrystallization aggregation of canavalin by dynamic light scattering,” *Biophys J*, vol. 57, no. 1, pp. 125–32, 1990.

[216] M. Skouri, M. Delsanti, J. P. Munch, B. Lorber, and R. Giege, “Dynamic light scattering studies of the aggregation of lysozyme under crystallization conditions,” *FEBS Lett*, vol. 295, no. 1-3, pp. 84–8, 1991.

[217] Y. Georgalis, E. B. Starikov, B. Hollenbach, R. Lurz, E. Scherzinger, W. Saenger, H. Lehrach, and E. E. Wanker, “Huntingtin aggregation monitored by dynamic light scattering,” *Proc Natl Acad Sci U S A*, vol. 95, no. 11, pp. 6118–21, 1998.

[218] B. Jachimska, M. Wasilewska, and Z. Adamczyk, “Characterization of globular protein solutions by dynamic light scattering, electrophoretic mobility, and viscosity measurements,” *Langmuir*, vol. 24, no. 13, pp. 6866–72, 2008.

[219] O. K. Gasymov and B. J. Glasgow, “Ans fluorescence: potential to augment the identification of the external binding sites of proteins,” *Biochim Biophys Acta*, vol. 1774, no. 3, pp. 403–11, 2007.

[220] D. J. Merkler, K. Srikumar, S. P. Marchese-Ragona, and F. C. Wedler, “Aggregation and thermo-inactivation of glutamine synthetase from an extreme thermophile, bacillus caldolyticus,” *Biochim Biophys Acta*, vol. 952, no. 1, pp. 101–14, 1988.

[221] E. N. Salgado, X. I. Ambroggio, J. D. Brodin, R. A. Lewis, B. Kuhlman, and F. A. Tezcan, "Metal templated design of protein interfaces," *Proc Natl Acad Sci U S A*, vol. 107, no. 5, pp. 1827–32, 2010.

[222] E. N. Salgado, R. A. Lewis, J. Faraone-Mennella, and F. A. Tezcan, "Metal-mediated self-assembly of protein superstructures: influence of secondary interactions on protein oligomerization and aggregation," *J Am Chem Soc*, vol. 130, no. 19, pp. 6082–4, 2008.

[223] Y. P. Pang, K. Xu, J. E. Yazal, and F. G. Prendergas, "Successful molecular dynamics simulation of the zinc-bound farnesyltransferase using the cationic dummy atom approach," *Protein Sci*, vol. 9, no. 10, pp. 1857–65, 2000.

[224] Y. Miller, B. Ma, and R. Nussinov, "Zinc ions promote alzheimer abeta aggregation via population shift of polymorphic states," *Proc Natl Acad Sci U S A*, vol. 107, no. 21, pp. 9490–5, 2010.

[225] W. T. Chen, Y. H. Liao, H. M. Yu, I. H. Cheng, and Y. R. Chen, "Distinct effects of zn2+, cu2+, fe3+, and al3+ on amyloid-beta stability, oligomerization, and aggregation: amyloid-beta destabilization promotes annular protofibril formation," *J Biol Chem*, vol. 286, no. 11, pp. 9646–56, 2011.

A APPENDIX

A.1 Electrophoresis Marker Legends

A.2 CHARMM Parameters

A.3 Vector Maps

A.4 Statistical Analysis of DLS Data

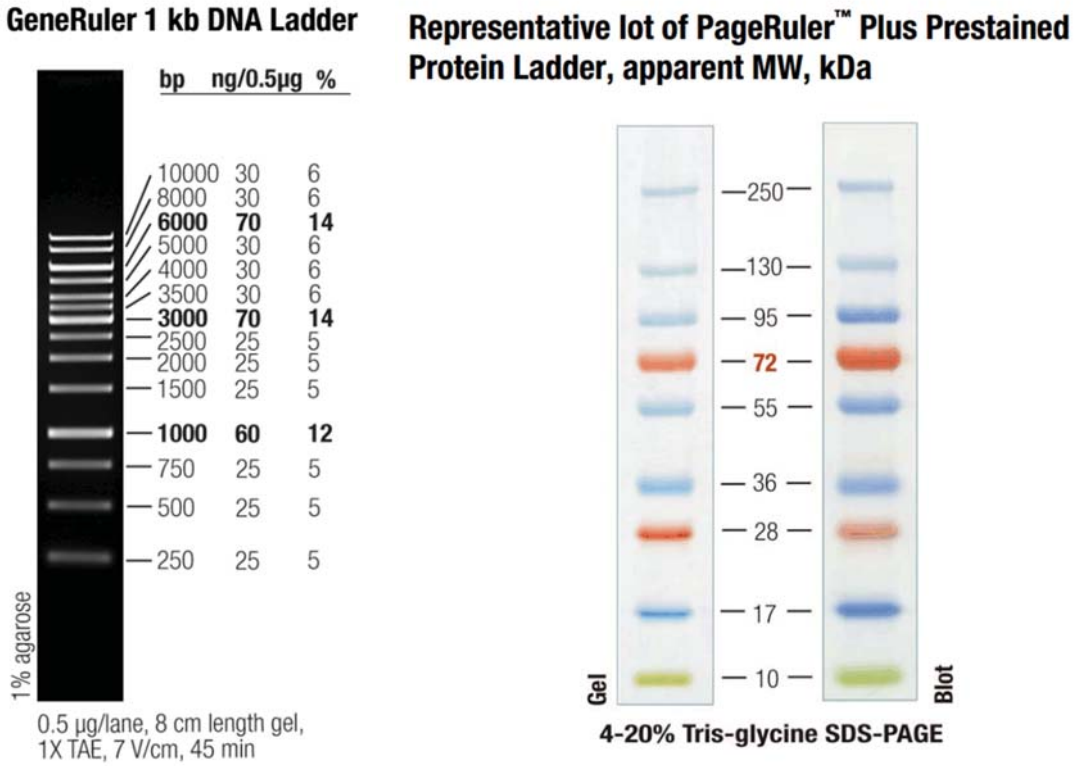

Figure A.1: DNA and Protein Molecular Weight Markers

```
!tributyrin
!
! Polar Head and glycerol backbone
GROUP                    !          ---------------------------------
ATOM C1   CTL2   -0.05 !          |                                 |
ATOM HA   HAL2    0.09 !          |                                 |
ATOM HB   HAL2    0.09 !    H2A---C12---H2B                         |
ATOM O11  OSL    -0.34 !          |                                 |
ATOM C11  CL      0.63 !          C11                               |
ATOM O12  OBL    -0.52 !         // \                              |
ATOM C12  CTL2   -0.08 !       O12  O11                            |
ATOM H2A  HAL2    0.09 !          |                                 |
ATOM H2B  HAL2    0.09 !       HX---C1---HY                       |
GROUP                    !          |                              |
ATOM C2   CTL1    0.04 !       HS---C2--------------              |
ATOM HS   HAL1    0.09 !            | beta1        |              |
ATOM O21  OSL    -0.34 !       O22  O21          theta3            |
ATOM C21  CL      0.63 !        \\ /  beta2        |               |
ATOM O22  OBL    -0.52 !         C21               |               |
ATOM C22  CTL2   -0.08 !         |   beta3         |               |
ATOM H2R  HAL2    0.09 !   H2R---C22---H2S        |               |
ATOM H2S  HAL2    0.09 !          |                |               |
GROUP                    !          |    beta4       |               |
ATOM C3   CTL2   -0.05 !          |           HX---C3---HY         |
ATOM HX   HAL2    0.09 !          |                |   gamma1      |
ATOM HY   HAL2    0.09 !          |           O32  O31             |
ATOM O31  OSL    -0.34 !          |            \\ /    gamma2      |
ATOM C31  CL      0.63 !          |             C31                |
ATOM O32  OBL    -0.52 !          |             |      gamma3      |
ATOM C32  CTL2   -0.08 !          |        H2X---C32---H2Y         |
ATOM H2X  HAL2    0.09 !          |                | gamma4        |
ATOM H2Y  HAL2    0.09 !          |                |               |
GROUP                    !          |                |               |
ATOM C13  CTL2   -0.18 !          |                |       H3A---C13---H3B
ATOM H3A  HAL2    0.09 !          |                |               |
ATOM H3B  HAL2    0.09 !          |                |               |
GROUP                    !          |                |               |
ATOM C14 CTL3    -0.27 !          |                |       H4A---C14--H4B
ATOM H4A HAL3     0.09 !          |                |               |
ATOM H4B HAL3     0.09 !          |                |             H4C
ATOM H4C HAL3     0.09 !          |                |
GROUP                    !          |                |
ATOM C23  CTL2   -0.18 !          |                |
ATOM H3R  HAL2    0.09 !   H3R ---C23---H3S       |
ATOM H3S  HAL2    0.09 !          |                |
```

```
GROUP                   !           |                 |
ATOM C24 CTL3    -0.27 !           |                 |
ATOM H4R HAL3     0.09 !   H4R---C24--H4S           |
ATOM H4S HAL3     0.09 !           |                 |
ATOM H4T HAL3     0.09 !         H4T                 |
GROUP                   !                           |
ATOM C33  CTL2   -0.18 !                           |
ATOM H3X  HAL2    0.09 !                  H3X ---C33---H3Y
ATOM H3Y  HAL2    0.09 !                           |
GROUP                   !                           |
ATOM C34 CTL3    -0.27 !                           |
ATOM H4X HAL3     0.09 !                    H4X---C34--H4Y
ATOM H4Y HAL3     0.09 !                           |
ATOM H4Z HAL3     0.09 !                         H4Z
!
! Glycerol Backbone
BOND  C1   HA        C1    HB       C1   C2        C1   O11
BOND  C2   HS        C2    C3       C2   O21
BOND  C3   HX        C3    HY       C3   O31
! Chain from C1
BOND  O11  C11
BOND  C11  C12
DOUBLE  C11  O12
BOND  C12  H2A       C12   H2B      C12  C13
BOND  C13  H3A       C13   H3B      C13  C14
BOND  C14  H4A       C14   H4B      C14  H4C
! Chain from C2
BOND  O21  C21
BOND  C21  C22
DOUBLE  C21  O22
BOND  C22  H2R       C22   H2S      C22  C23
BOND  C23  H3R       C23   H3S      C23  C24
BOND  C24  H4R       C24   H4S      C24  H4T
! Chain From C3
BOND  O31  C31
BOND  C31  C32
DOUBLE  C31  O32
BOND  C32  H2X       C32   H2Y      C32  C33
BOND  C33  H3X       C33   H3Y      C33  C34
BOND  C34  H4X       C34   H4Y      C34  H4Z

!IC table from IC generate, geometry is guessed
IC C2   O11  *C1  HA   1.5458  109.83  118.28  108.95   1.1134
IC HA   O11  *C1  HB   1.1134  108.95  118.31  112.84   1.1153
IC O11  C1   C2   O21  1.4279  109.83 -169.38  109.43   1.4399
```

```
IC O21  C1   *C2  C3   1.4399  109.43 -120.76  111.51   1.5508
IC C3   C1   *C2  HS   1.5508  111.51 -116.85  107.06   1.1136
IC C1   C2   O21  C21  1.5458  109.43  123.06  117.42   1.3213
IC C2   O21  C21  C22  1.4399  117.42  163.64  109.74   1.5303
IC C22  O21  *C21 O22  1.5303  109.74 -176.86  126.01   1.2239
IC O21  C21  C22  C23  1.3213  109.74 -113.72  109.53   1.5428
IC C23  C21  *C22 H2R  1.5428  109.53  120.69  110.65   1.1107
IC H2R  C21  *C22 H2S  1.1107  110.65  118.78  108.83   1.1092

IC C3   C2   C1   O11  1.5458  111.51 -171.21  111.32   1.4465
IC O11  C2   *C1  HA   1.4465  111.32 -126.25  108.41   1.1159
IC HA   C2   *C1  HB   1.1159  108.41 -114.81  107.61   1.1132
IC C2   C1   O11  C11  1.5508  111.32 -117.52  117.59   1.3317
IC C1   O11  C11  C12  1.4465  117.59 -177.93  109.00   1.5333
IC C12  O11  *C11 O12  1.5333  109.00 -179.19  126.42   1.2107
IC O11  C11  C12  C13  1.3317  109.00 -176.00  111.40   1.5440
IC C13  C11  *C12 H2A  1.5440  111.40  120.79  108.80   1.1093
IC H2A  C11  *C12 H2B  1.1093  108.80  118.57  108.70   1.1101

IC C1   C2   C3   O31  1.5458  111.51 -171.21  111.32   1.4465
IC O31  C2   *C3  HX   1.4465  111.32 -126.25  108.41   1.1159
IC HX   C2   *C3  HY   1.1159  108.41 -114.81  107.61   1.1132
IC C2   C3   O31  C31  1.5508  111.32 -117.52  117.59   1.3317
IC C3   O31  C31  C32  1.4465  117.59 -177.93  109.00   1.5333
IC C32  O31  *C31 O32  1.5333  109.00 -179.19  126.42   1.2107
IC O31  C31  C32  C33  1.3317  109.00 -176.00  111.40   1.5440
IC C33  C31  *C32 H2X  1.5440  111.40  120.79  108.80   1.1093
IC H2X  C31  *C32 H2Y  1.1093  108.80  118.57  108.70   1.1101

IC C11  C12  C13  C14  1.5303  109.53  168.08  112.83   1.5322
IC C14  C12  *C13 H3A  1.5322  112.83 -119.08  109.34   1.1146
IC H3A  C12  *C13 H3B  1.1146  109.34 -118.17  109.92   1.1133
!IC C12  C13  C14  C15  1.5428  112.83  179.85  112.06   1.5333
!IC C15  C13  *C14 H4A  1.5333  112.06  120.31  108.77   1.1150
IC H4A  C13  *C14 H4B  1.1150  108.77  117.76  110.16   1.1129
IC H4A C13 *C11 H4C 1.1113  110.37  120.07  110.59   1.1111
IC C21  C22  C23  C24  1.5303  109.53  168.08  112.83   1.5322
IC C24  C22  *C23 H3R  1.5322  112.83 -119.08  109.34   1.1146
IC H3R  C22  *C23 H3S  1.1146  109.34 -118.17  109.92   1.1133
!IC C22  C23  C24  C25  1.5428  112.83  179.85  112.06   1.5333
!IC C25  C23  *C24 H4R  1.5333  112.06  120.31  108.77   1.1150
IC H4R  C23  *C24 H4S  1.1150  108.77  117.76  110.16   1.1129
IC H4R C23 *C21 H4T 1.1113  110.37  120.07  110.59   1.1111
IC C31  C32  C33  C34  1.5333  111.40  179.66  112.54   1.5345
IC C34  C32  *C33 H3X  1.5345  112.54 -121.73  109.66   1.1133
IC H3X  C32  *C33 H3Y  1.1133  109.66 -117.45  109.37   1.1151
!IC C32  C33  C34  C35  1.5440  112.54  178.40  112.54   1.5346
!IC C35  C33  *C34 H4X  1.5346  112.54  121.75  110.03   1.1131
IC H4X  C33  *C34 H4Y  1.1131  110.03  117.75  108.84   1.1139
IC H4X C33 *C31 H4Z 1.1113  110.07  120.81  110.70   1.1116
```

Figure A.2: CHARMM Parameters for Tributyrin

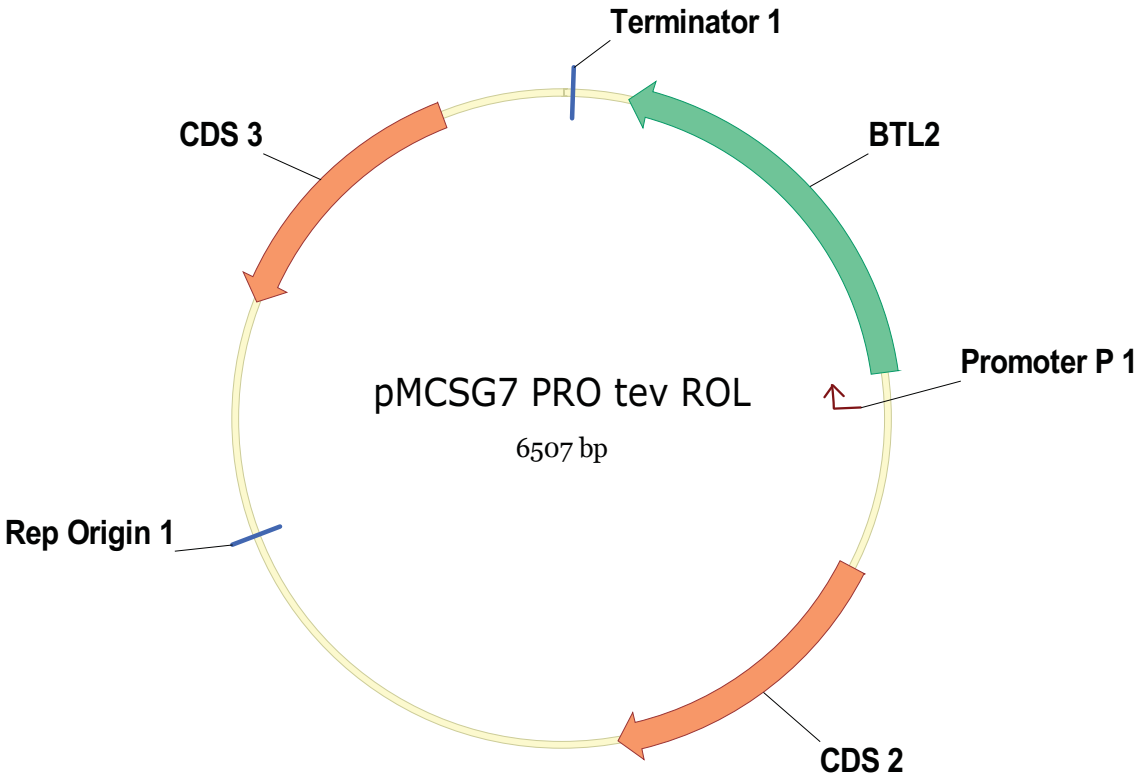

Figure A.3: Cloning Vector Map

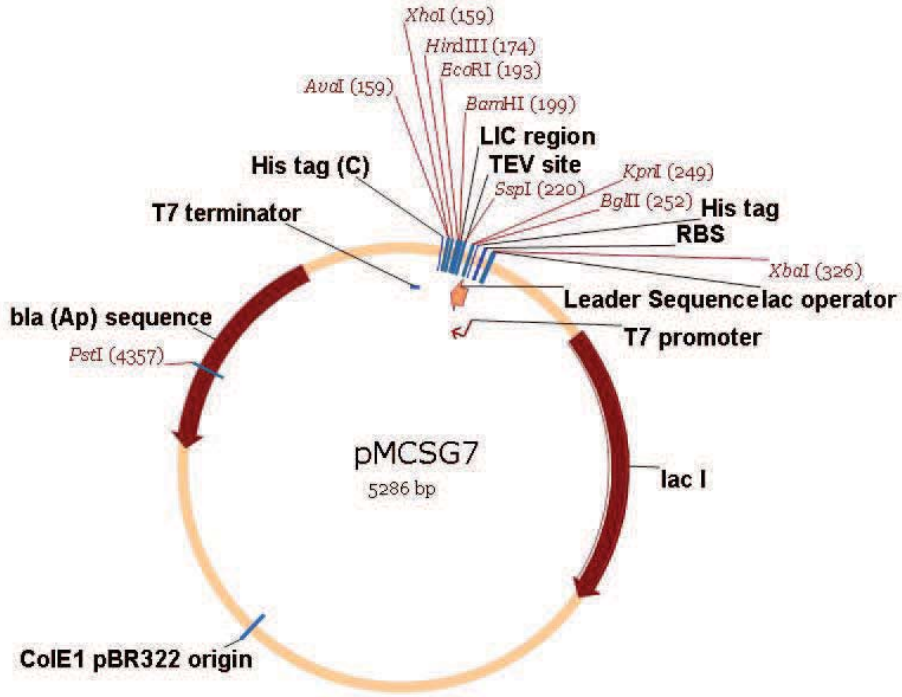

Figure A.4: Expression Vector Map

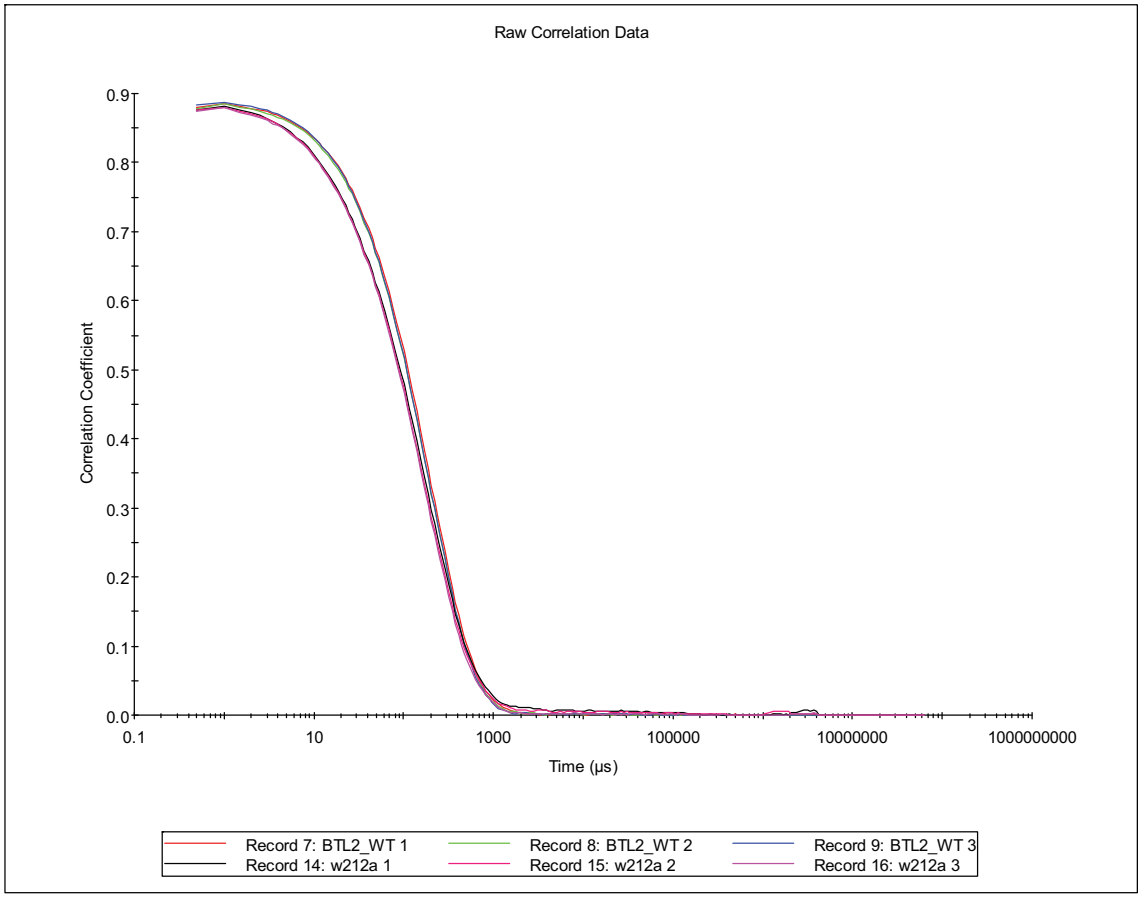

Figure A.5: Correlation Data

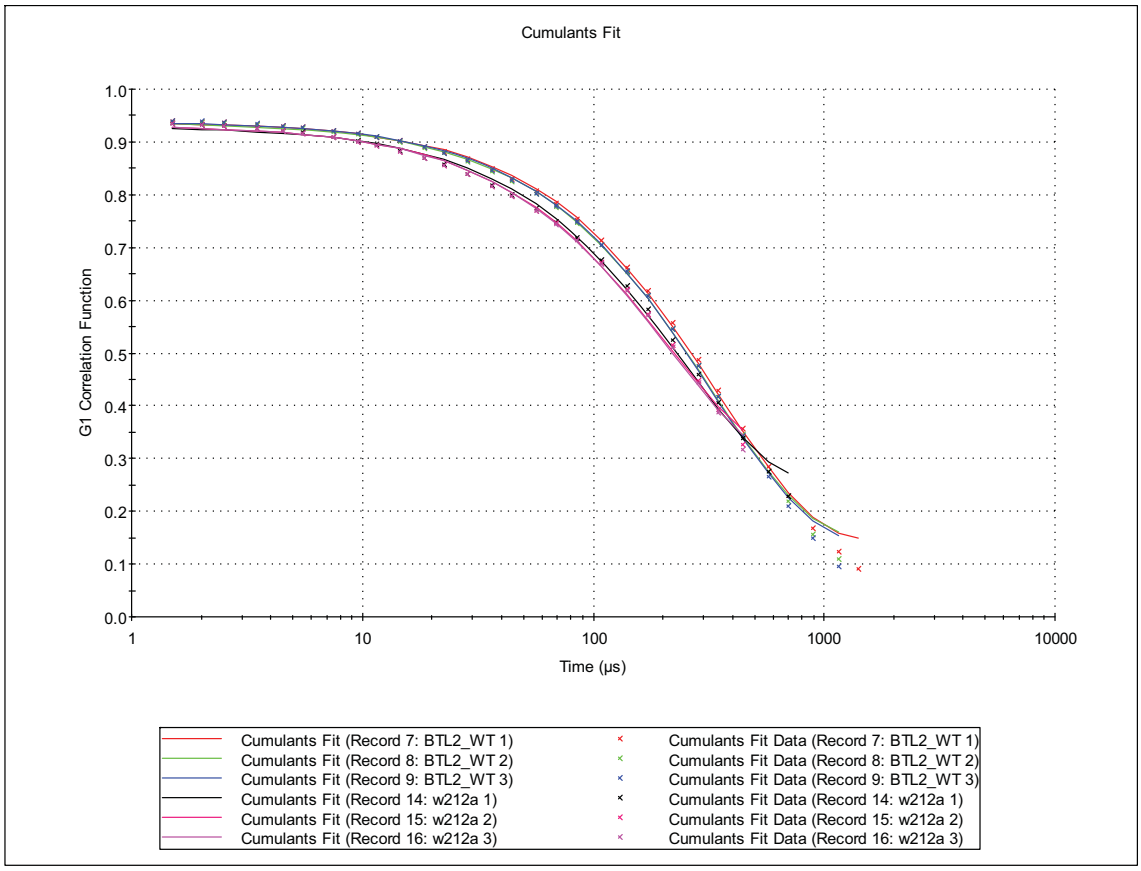

Figure A.6: Cumulants-fit Data